

Prediction of milk yield using visual images of cows through deep learning

By

LAWRENCE JEMBERE

A dissertation submitted in fulfilment of the requirements for the degree of

Master of Science in Agriculture (Animal Science)

in the

School of Agricultural, Earth and Environmental Sciences

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Scottsville, Pietermaritzburg

South Africa




Supervisor: Professor Michael Chimonyo

September 2022

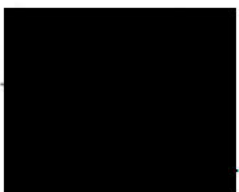
EXAMINER'S COPY

DECLARATION

I, Lawrence Jembere, declare that this thesis entitled “Prediction of milk yield using cow’s visual images through deep learning” is originally my work, apart from where stated through citations or quotes. This dissertation was conducted under the supervision of Professor Michael Chimonyo and has not been submitted to any other University other than the University of KwaZulu-Natal.

Jembere L (author)...  Date...23 November 2022.....

As the candidate’s supervisor, I agree to the submission of this thesis

Prof. M. Chimonyo (supervisor)...  Date: 5 December 2022

Thesis output

Conference abstract

Jembere L and Chimonyo M (2021). Prediction of milk yield using cows' visual images through deep learning. College of Agriculture, Engineering and Science Research and Innovation Symposium, 10 December 2021. University of KwaZulu-Natal, Pietermaritzburg.

Jembere L (author)... 

Acknowledgements

First, my deep and sincere gratitude goes to my supervisor, Professor Michael Chimonyo, for his support, constructive criticism, guidance and financial assistance during the ordeal times on my study.

I also thank the dairy farmers who allowed me to collect data at their farms and offered accommodation during data collection. These farms are Bloekomlaan dairy farm, Sand water dairies, Glen herd farm, Milagro farm and Bo Plaas farm. I especially thank the following individuals for going out of their way to assist me with data collection: Chris Do Plooy, Illa Hugo, Madaleen Lombard, Izak van der Merwe, Clark Brink and Andrew Masterson.

I am ineffably grateful for my family's support and motivation and for always being there for me. Special thanks go to my brother Dr E Jembere for being there for me, not only as a brother but also as a lecturer, especially in the field of computer science. I also thank Prof S Viriri for guiding me through computer vision, deep learning and artificial intelligence concepts. Not forgetting all my friends who were there for me in various ways.

Finally, I thank God for all he has done for me, where he took me from, where I am and where he is taking me. I am profusely grateful for all the people he has allowed me to cross paths with.

Abstract

Prediction of milk yield using visual images of cows through deep learning

By

Lawrence Jembere

The broad objective of the study was to determine, through deep learning, the predictability of milk yield from a cow's image data. The data size of 1238 image pairs (the side-view images and the rear-view images) from 743 Holstein cows within their first or second parity and the cows' corresponding first lactation 305-day milk yield values were used to train a deep learning model. The data was first split into the training and testing data at the ratio of 80:20, respectively. The training data was then augmented four times more, then again split into training and validation data at the ratio of 80:20, respectively.

Three principal analyses were done, i.e. the prediction of milk yield using rear-view images only, the prediction of milk yield using the side-view images only and the prediction of milk yield using a merge of the side-view and rear-view images (the combined-view images). In all three analyses, poor predictions were observed, i.e. R^2 values of 0.32 for the model using the side-view image, 0.30 for the model using the rear-view images and 0.38 for the model using combined side and rear images. The mean absolute errors were 1146.4 kg, 1148.3 kg and 1112.9 kg for the side-view, the rear-view and the combined-view models, respectively. The root mean square error values were 1460.7 kg, 1480.5 kg and 1401.2 kg and the mean absolute error percentages were 17.6, 17.3 and 17.0 % for the side-view, rear-view and combined-view models, respectively.

Hypotheses tests were also done to check whether there was any difference between these three prediction models. There was no significant difference in performance between all the prediction models ($p > 0.05$), i.e. the side-view model, the rear-view model and the combined-view model. It was concluded that predicting 305-day milk yield of Holstein cows using either view has the same level of accuracy and no additional benefits are derived from using both the rear and the side views.

Keywords: Computer vision; deep learning; linear conformation traits; 305-day milk yield; side-view images; rear-view images; combined-view images; Holstein cows.

List of abbreviations

AI	Artificial intelligence
ANG	Angularity
ANN	Artificial Neural Network
BCS	Body condition score
BLUP	Best Linear Unbiased Prediction
CNN	Convolutional Neural Network
CSV	Comma-separated values
CV	Computer Vision
DL	deep learning
DNN	deep neural network
GBLUP	Genomic Best Linear Unbiased Prediction
ICAR	International Committee for Animal Recording
ID	Identity
LR	Learning rate
LTT	Linear-type traits/linear conformation traits
MAE	Mean absolute error
MAPE	Mean absolute percentage error
ML	Machine learning

MSE	Mean square error
MY	Milk yield
QTL	Quantitative trait loci
ReLU	Rectified Linear Unit
RMSE	Root mean square error
RUH	Rear udder height
RUW	Rear udder width
RW	Rump width
R^2	Coefficient of determination
UD	Udder depth

Dedication

This dissertation is devoted to my parents, five siblings and their spouses.

Table of Contents

DECLARATION	ii
Thesis output	iii
Acknowledgements	iv
Abstract	v
List of abbreviations	vi
Dedication	viii
Table of Contents	ix
List of Tables	xii
List of Figures	xiv
1. CHAPTER ONE: Introduction	1
1.1. Background	1
1.2. Justification	3
1.3. Objectives.....	4
1.4. Hypotheses.....	4
1.5. References.....	5
2. CHAPTER TWO: Literature review	7
2.1. Introduction	7
2.2. The South African dairy industry.....	7
2.3. Milk yield prediction	8
2.3.1. Components used to predetermine milk yield	9
2.3.2. Correlation between linear-type traits and milk yield	17
2.3.3. Milk yield prediction using linear-type traits	18
2.4. Computer vision deep learning	19
2.4.1. Convolutional neural networks.....	20
2.4.2. Challenges of computer vision deep learning.....	21
2.4.3. Uses of computer vision in animal production	22
2.4.4. Advantages of using computer vision deep learning in predicting milk yield	26
2.5. Data augmentation and the augmentation types.....	26
2.6. Summary	27
2.7. References.....	28
3. CHAPTER THREE: Predicting milk yield using a cow’s rear-view image through computer vision deep learning	36
Abstract.....	36
3.1. Introduction	36
3.2. Materials and Methods.....	37

3.2.1.	Description of the study sites.....	37
3.2.2.	Study variables	38
3.2.3.	The data	38
3.2.4.	Data pre-processing	40
3.2.5.	Data analysis	43
3.2.6.	Description of the data	50
3.3.	Results and Discussion	51
3.3.1.	Effects of splitting the test data after augmentation.....	53
3.4.	Conclusions	56
3.5.	References.....	58
4.	CHAPTER FOUR: Predicting milk yield using a cow’s side-view image through computer vision deep learning	60
	Abstract.....	60
4.1.	Introduction	60
4.2.	Hypothesis.....	61
4.3.	Materials and Methods.....	61
4.3.1.	Description of the study sites.....	61
4.3.2.	The study variables	61
4.3.3.	The data	62
4.3.4.	Data pre-processing	62
4.3.5.	Data analysis	62
4.3.6.	Description of the data	67
4.4.	Results and Discussion	67
4.4.1.	Side-view model vs rear-view model	69
4.4.2.	Effects of splitting the test data after augmentation.....	70
4.5.	Conclusions	72
4.6.	References.....	73
5.	CHAPTER FIVE: Predicting milk yield using a combination of cows’ side-view and rear-view images through computer vision deep learning	75
	Abstract.....	75
5.1.	Introduction	75
5.2.	Hypotheses.....	76
5.3.	Materials and Methods.....	76
5.3.1.	The study sites.....	76
5.3.2.	The study variables	77
5.3.3.	The data	77

5.3.4.	Data pre-processing	77
5.3.5.	Data analysis	78
5.3.6.	Description of the data	85
5.4.	Results and Discussion	85
5.4.1.	Effects of splitting the test data after augmentation.....	88
5.5.	Conclusions	90
5.6.	References.....	91
6.	CHAPTER SIX: General discussion, Conclusions and Recommendations.....	101
6.1.	General discussion	101
6.2.	Conclusions	102
6.3.	Recommendations	102
6.4.	References.....	104
Appendix 1: Data collection approval letter from The Glen farm		105
.....		105
Appendix 2: Data collection approval letter from Sandwater dairy farm		106
Appendix 3: Data collection approval letter from Milagro farm		107
Appendix 4: Data collection approval letter from Bloekomlaan dairy farm.....		108

List of Tables

Table 3.1: Description of the data	39
Table 3.2: Rear-view learning rate comparison results at 38 epochs, with LR scheduler and no dropout.	46
Table 3.3: Comparison between a model with a learning rate scheduler and one with no LR scheduler.	48
Table 3.4: Effects of dropouts (dropout rate=0.4) on model learning.....	49
Table 3.5: Optimum parameter used for the analyses.	50
Table 3.6: Description of the testing and the training and validation data in terms of sample size, average, maximum, minimum and standard deviation of the 305-day milk yield.....	51
Table 3.7: Training, validation and testing results for the rear-view model with no dropouts, learning rate=0.0006, 38 epochs and a learning rate scheduler.	52
Table 3.8: Training, validation and testing results for the rear-view model with the test data split after augmentation, no dropouts, learning rate=0.0006, 38 epochs and a learning rate scheduler.	54
Table 4.1: Side-view learning rate comparison results at 35 epochs, with LR scheduler and no dropout.	64
Table 4.2: Comparison between a side-view model with a learning rate scheduler and one with no LR scheduler.	65
Table 4.3: effects of dropouts (dropout rate=0.4) on model learning.	66
Table 4.4: Optimum parameter used for the final side-view analyses.	67
Table 4.5: Training, validation, and testing results for the side-view model with no dropouts, a learning rate of 0.0006, 38 epochs and a learning rate scheduler.....	68
Table 4.6: T-test: Two-tailed, paired t-test between the side-view model and the rear-view model.....	69
Table 4.7: Training, validation and testing results for the side-view model with the test data split after augmentation, no dropouts, no learning rate scheduler, initial learning rate= 0.0005 and 35 epochs.....	70
Table 5.1: Combined-view learning rate comparison results at 50 epochs, with no LR scheduler and no dropout.	82
Table 5.2: Comparison between a combined-view model with a learning rate scheduler and one with no LR scheduler.	83
Table 5.3: effects of dropouts (dropout rate=0.4) on the combined-view model learning....	84
Table 5.4: Optimum parameter used for the final side-view analyses.	85
Table 5.5: Training, validation, and testing results for the combined-view model with no dropouts, a learning rate of 0.0006, 50 epochs and no learning rate scheduler.....	86
Table 5.6: Two-tailed paired t-test between the rear-view and the combined-view models.	86

Table 5.7: Two-tailed paired t-test between the side-view and combined-view models.	87
Table 5.8: Training, validation and testing results for the combined-view model with the test data split after augmentation, no dropouts, no learning rate scheduler, initial learning rate=0.0006 and 50 epochs.....	88

List of Figures

Figure 2.1: Cow images showing various rear udder widths. Source: ICAR (2015).....	13
Figure 2.2: Cow images showing various udder depths. Source: ICAR (2015).....	13
Figure 2.3: Cow images showing various rear udder heights. Source ICAR (2015).....	13
Figure 2.4: Cow images showing various fore udder attachments. Source: ICAR (2015) ..	14
Figure 2.5: Cow images showing various angularities. Source: ICAR (2015).....	15
Figure 2.6: Cow images showing various body depths. Source: ICAR (2015).....	15
Figure 2.7: Cow images showing how stature is measured. Image A shows the rear view of how stature is measured and Image B shows the point of measurement. Source: ICAR (2015).	16
Figure 2.8: : Cow images showing various rump width. Source: ICAR (2015).....	17
Figure 2.9: Relationship between computer vision, deep learning, machine learning and Artificial Intelligence. (Johnson 2019)	20
Figure 2.10: A schematic representation of the building blocks of a convolutional neural network.....	21
Figure 2.11: Representative images with data augmentation techniques, such as flipping (b), scaling(c), rotating (d), stretching (e), noise (f), padding(g), cropping (h), and shearing (i).27	
Figure 3.1: Image before shedding out undesired cows(left); image after removing undesired cows(right)	42
Figure 3.2: Rear-view learning rate comparison results at 38 epochs, with LR scheduler and no dropout.	47
Figure 3.3: Rear-view learning curves for the training and validation.	49
Figure 3.4: 305-day milk yield frequency distribution for the training and validation data: 3.4a.(left) and the testing data: 3.4b. (right).....	51
Figure 3.5: Scatter plot of the predicted vs the actual milk yield values from the rear-view model where the test data is split after augmentation.	55
Figure 3.6: Scatter plot of the predicted vs the actual milk yield values from the rear-view model where the test data is split before augmentation.	56
Figure 4.1: Side-view learning rate comparison results at 38 epochs, with LR scheduler and no dropout.	65
Figure 4.2: Side-view learning curves for the training and validation.....	66
Figure 4.3: Scatter plot of the predicted vs the actual milk yield values from the side-view model where the test data is split after augmentation.	71
Figure 4.4: Scatter plot of the predicted vs the actual milk yield values from the side-view model where the test data is split before augmentation.	72
Figure 5.1: Combined-view image with disproportionate side and rear view images.	79

Figure 5.2: Combined images with the same image proportion for the side-view and rear-view	80
Figure 5.3: Combined-view learning rate comparison results at 50 epochs, with no LR scheduler and no dropout.	83
Figure 5.4: Combined-view learning curves for the training and validation.	84
Figure 5.5: Scatter plot of the predicted vs the actual milk yield values from the combined-view model where the test data is split after augmentation.....	89
Figure 5.6: Scatter plot of the predicted vs the actual milk yield values from the combined-view model where the test data is split before augmentation.....	90

CHAPTER ONE: Introduction

1.1. Background

Dairy production plays a critical role in employment creation and food security. In South Africa, the dairy industry is the fifth largest agricultural sector and one of the highest employment creation sectors (Ogundeji et al., 2021). Due to the economic disturbances from the onset of COVID-19, dairy production costs have sharply increased over the past one and half years, primarily maize and soya prices (FAO, 2021; Milk Producers Organization, 2020). Based on the FAO report for the year 2020, South Africa's milk production dropped by about 5% in 2019 (FAO, 2021). This has resulted in job losses and the need to cull low-producing cows to have a smaller herd of high-producing cows.

Now more than ever, dairy farmers desire to predict each cow's milk yield, especially when they are still heifers. Such predictions would aid farmers in making better cow replacement decisions, financial planning, and detecting deviating yield patterns that may indicate mastitis. Deep learning models are a new approach to predicting milk yield based on visual images of cows.

Various linear-type traits are correlated with milk yield (MY), e.g. angularity (ANG), rump width (RW) and most udder traits (Campos et al., 2015; Harris, 2015; Khan & Khan, 2016). Since these traits can be seen earlier in a cow's life, they are often used to predict milk yield and quality. The variability in the results of these correlations is liberal. For example, Campos et al. (2015) found a phenotypic correlation of 0.19 between rear udder width and milk yield, whereas Khan & Khan (2016) found the correlation between milk yield and rear udder width to be 0.54. Such variability can be bewildering in making predictions, especially when establishing trait weights in selection indices. Such variations can be imputed to differences in measurement methods as some use actual measurements of the linear-type traits (Khan & Khan, 2016), and some use visual appraisal (Campos et al., 2015). Khan & Khan (2016) attributed variations to differences in visual judgment for traits such as angularity that can only be determined through visual inspection. Therefore, it is important to consider using computer vision deep learning as a predictive tool to increase the accuracy of prediction.

Deep learning is a machine learning method based on artificial neural networks (ANN) with representation learning (Hordri et al., 2017). Deep learning through convolutional neural networks (CNN) makes it possible to use visual data for prediction; this is called computer vision. A convolutional neural network is a subcategory of ANN composed of numerous

building blocks, such as convolution layers, pooling layers, and fully connected layers (Patil & Rane, 2021). Convolutional neural networks can extract features on an image, put weight on each feature based on the desired output, and then make predictions. The model automatically and adaptively trains and betters itself through backpropagation (Patel et al., 2010). It is unclear whether milk yield can be accurately predicted using machine learning.

Linear-type traits (LTT) or conformation traits are objective descriptions of an animal's body (Harris, 2015). They describe the degree of a feature, not the desirability (ICAR, 2015). Linear-type traits are visible and can be seen earlier in a cow's life, making them salient indicators of production traits with which they are correlated. The side and rear views of cows can provide images of the udder, rump, and angularity. Unlike other prediction methods such as genomic selection, taking images does not require special skills or expert knowledge. The images can easily be taken using a cellphone, making milk yield prediction easy and convenient.

The use of computer vision deep learning in predicting milk yield is currently low, although it has a huge potential in predicting milk yield (Ozkaya, 2015). One of the challenges with using computer vision deep learning for predicting milk yield is that it does not directly report the specific features it used to make predictions and the weights assigned to each feature. There may be a need for feature silencing to get more specified information on the actual traits used for prediction and their contribution to predicting milk yield. Capturing all the features necessary for milk predictions in one picture may be challenging.

In 2020, raw milk production in South Africa was 0.31 % lower than in 2019. This is due to the recent sharp increase in maize and soya prices pertinent to making dairy feed (South African Milk Processors' Organization, 2020). High feed prices and the reduction in personnel have resulted in the need to cull down on low-producing cows. Therefore, every dairy farmer desires to predetermine the amount of milk each cow in his herd will produce. Pre-exposure to this information equips the farmer with better decisions about culling, feeding, selling, and even mating the cows. Farmers, however, cannot use techniques such as genomic selection, with higher prediction accuracy, on all the animals in the herd because it is expensive, and the service is less available at the national level. Farmers often resort to using linear-type traits correlated with milk yield to make milk yield predictions. This is because personnel to perform this task is usually available. However, it cannot be ignored that the process of trait measuring and trait weighting in selection indices is laborious, time-consuming and requires some level of expertise. Also, the weighting of traits in the selection indices is dependent on the literature used. The question then is, can computer vision deep learning be deployed as a prediction tool for milk yield using linear-type traits? Deep learning allows the machine to do the

computations and assign necessary weights to the traits for prediction, reducing possible bias and labour in the prediction analysis.

1.2. Justification

Models to predict milk production enable farmers to predict their herd's productive capacity and future earnings (Ehrlich, 2011; Liseune et al., 2020). Predicting lactation details facilitates the selection process as it identifies the most productive females and superior bulls based on the analysis of the total productivity of their offspring (Liseune et al., 2020). Moreover, farmers can assess the required feed intake, plant utilization and energy and protein requirement by estimating the expected milk yield for each cow, allowing them to evaluate their costs (Zhang et al., 2018; Liseune et al., 2020). Comparing a cow's predicted versus actual milk yield enables the detection of diseases and other factors of cow health, facilitating improved animal monitoring systems (Liseune et al., 2020). Early detection of unproductive cows also supports informed culling decisions (Njubi et al., 2010).

Conformation traits are often used to predict milk yield as they can be measured earlier in a cow's life (Getu & Misganaw, 2015). Available prediction models based on linear conformation traits are susceptible to bias and inaccuracies due to variations in human perception (Khan & Khan, 2016). For example, there is conflict in the literature on whether correlations of rump width with milk yield are essential for predictions. Several studies found a weak correlation between milk yield and rump width (Harris et al., 1992; Klassen et al., 1992; Brotherstone, 1994), while others reported a moderate phenotypic correlation between milk yield and rump width (Harris, 2015). There is, therefore, a need to provide an automated way of predicting milk yield. The prediction requires computer scientists to develop software that farmers can use for predictions using image data.

The conventional process of computing/predicting milk yield from these linear-type traits is cumbersome and labour-intensive. Image data also broadens the scope of data acceptable for analysis since image data is permissible in computer vision machine learning.

Machines can handle and compute more data faster than humans (Kao & Venkatachalam, 2021). Given labelled data, computers can quickly search for relationships between the data and the labels.

Genomic selection can also be performed using imagery data, making it cost-effective to select superior cows. The image data can be combined with genomic selection, trait heritabilities and pedigree information to estimate breeding values and performance.

1.3. Objectives

The broad objective of the study was to determine, through deep learning, the predictability of milk yield from a cow's image data. The specific objectives were to:

1. Determine the predictability of 305-day milk yield using a cow's rear-view image.
2. Determine the predictability of 305-day milk yield using a cow's side-view image.
3. Determine the predictability of 305-day milk yield using a combination of a cow's side-view and rear-view image.

1.4. Hypotheses

The hypotheses tested in this study were:

1. Milk yield prediction is the same when using either the rear-view or the side-view images of cows.
2. milk yield prediction is the same when using either the combined-view or the rear-view images of cows.
3. Milk yield prediction is the same when using either the combined-view or the side-view images of cows.

1.5. References

- Brotherstone, S. (1994). Genetic and phenotypic correlations between linear type traits and production traits in Holstein-Friesian dairy cattle. *Animal Science*, 59(2), 183-187. doi:10.1017/S0003356100007662.
- Campos, R. V., Cobuci, J. A., Kern, E. L., Costa, C. N., & McManus, C. M. (2015). Genetic Parameters for Linear Type Traits and Milk, Fat, and Protein Production in Holstein Cows in Brazil. *Asian-Australasian Journal of Animal Sciences*, 28(4), 476-484. <https://doi.org/10.5713/ajas.14.0288>.
- Ehrlich, J. L. (2011). Quantifying shape of lactation curves , and benchmark curves for common dairy breeds and parities. *The Bovine Practitioner*, 45(1), 88–96. DOI: <https://doi.org/10.21423/bovine-vol45no1p88-95>
- FAO (2021) ‘Dairy market review, Food and Agriculture Organisation of the United Nations’, Food and Agriculture Organization of the United Nations, (April), pp. 1–13. Retrieved February 4, 2022 from the FAO website <https://www.fao.org/3/cb4230en/cb4230en.pdf>.
- Getu, A. and Misganaw, G. (2015) The Role of Conformational Traits on Dairy Cattle Production and Their Longevities. *Open Access Library Journal*, 2, 1-9. doi:10.4236/oalib.1101342.
- Grzesiak, W., Wojcik, J., & Binerowska, B. (2003). Prediction of 305-day first lactation milk yield in cows with selected regression models. *Archives Animal Breeding*, 46(3), 215–226. <https://doi.org/10.5194/aab-46-213-2003>
- Harris, B.L., A.E. Freeman and E. Metzger. 1992. Genetic and phenotypic parameters for type and production in Guernsey dairy cows. *J. Dairy Sci.* 75:1147-1153.
- Harris, R. (2015). Phenotypic correlations between linear type conformation traits, production and fertility in a once-a-day milked dairy cattle herd. Thesis retrieved 5 May, 2021 from the Massey university website: https://mro.massey.ac.nz/bitstream/handle/10179/7430/02_whole.pdf?sequence=2&isAllowed=y.
- Hordri, N. F., Samar, A., Yuhaniz, S. S., & Shamsuddin, S. M. (2017). A systematic literature review on features of deep learning in big data analytics. *International Journal of Advances in Soft Computing and Its Applications*, 9(1), 32–49. retrieved 20 August, 2021 from the Academia website: https://www.academia.edu/33049709/A_Systematic_Literature_Review_on_Features_of_Deep_Learning_in_Big_Data_Analytics.

- International Committee of Animal Recording. (2015). 1 . Conformation Recording of Dairy Cattle. June, 1–42. retrieved 1 August, 2021 from the ICAR website: <https://www.icar.org/wp-content/uploads/2015/08/Conformation-Recording-CR-WG.pdf>
- Kao, Y. F. and Venkatachalam, R. (2021) ‘Human and Machine Learning’, Computational Economics, 57(3), pp. 889–909. doi: 10.1007/s10614-018-9803-z.
- Khan, M. S. M. A., & Khan, M. S. M. A. (2016). Genetic and phenotypic correlations between linear type traits and milk yield in Sahiwal cows. Pakistan Journal of Agricultural Sciences, 53(2), 483–489. [doi.org: 10.21162/PAKJAS/16.3369](https://doi.org/10.21162/PAKJAS/16.3369)
- Klassen DJ, Monardes HG, Jairath L, Cue RI, Hayes JF. Genetic correlations between lifetime production and linearized type in Canadian Holsteins. J Dairy Sci. 1992 Aug;75(8):2272-82. doi: 10.3168/jds.S0022-0302(92)77988-0. PMID: 1401377.
- Liseune, A., Salamone, M., Van den Poel, D., Van Ranst, B., & Hostens, M. (2020). Leveraging latent representations for milk yield prediction and interpolation using deep learning. Computers and Electronics in Agriculture, 175(July), 105600. <https://doi.org/10.1016/j.compag.2020.105600>.
- Njubi, D. M., Wakhungu, J. W., & Badamana, M. S. (2010). Use of test-day records to predict first lactation 305-day milk yield using artificial neural network in Kenyan Holstein-Friesian dairy cows. Tropical Animal Health and Production, 42(4), 639–644. Doi: 10.1007/s11250-009-9468-7
- Ozkaya, S. (2015). Prediction possibility of milk yield from udder measurements using digital image analysis on Holstein cows. Indian Journal of Animal Research, 49(3), 388–391. Doi: 10.5958/0976-0555.2015.00050.3
- Patel, K., Bancroft, N., Drucker, S. M., Fogarty, J., Ko, A. J., & Landay, J. A. (2010). Gestalt: Integrated support for implementation and analysis in machine learning. UIST 2010 - 23rd ACM Symposium on User Interface Software and Technology, 37–46. Doi: 10.1145/1866029.1866038
- Patil, A., & Rane, M. (2021). Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition. Smart Innovation, Systems and Technologies, 195, 21–30. https://doi.org/10.1007/978-981-15-7078-0_3
- South African milk processors’ organization. (2020). Summary of Key Market Signals for the Dairy Industry , November 2020 Edition. Annual Report, November 2020. Retrieved 6 March 2021 from the South African milk processors’ organization website: <https://sampro.co.za/wp-content/uploads/2020/11/November-2020-edition-of-Summary-of-Key-Market-Signals-for-the-dairy-industry.pdf>
- Zhang, F., Shine, P., Upton, J., Shaloo, L., & Murphy, M. D. (2018). A review of milk production forecasting models: Past and future methods. Dairy Farming: Operations Management, Animal Welfare and Milk Production, January 2019, 13–61.
- Zink, V., Zavadilová, L., Lassen, J., Štípková, M., Vacek, M., & Štolc, L. (2014). Analyses of genetic relationships between linear type traits, fat-to-protein ratio, milk production traits, and somatic cell count in first-parity Czech Holstein cows. Czech Journal of Animal Science, 59(12), 539–547. <https://doi.org/10.17221/7793-cjas>

CHAPTER TWO: Literature review

2.1. Introduction

This chapter reviews the South African dairy industry, components used to predict milk yield, computer vision deep learning, and computer vision in animal production.

2.2. The South African dairy industry

The dairy industry ranks number five among the South African largest agricultural sectors and is one of the highest employment creation sectors (Ogundeji et al., 2021). It also contributes significantly to the country's food security. According to the South African Weather Service, the Western Cape, Eastern Cape, and KwaZulu-Natal receive up to five times more precipitation than South Africa's north-western and arid regions. Therefore, the higher rainfall supports the desired pasture-based grazing system. This explains why these three provinces contribute more than 70 % of South Africa's milk.

South Africa has about 1.2 million dairy cows, and the Major breeds are Holstein, Jersey and Ayrshire (Waal & Blom, 2020). Holstein is the most popular dairy breed, constituting approximately 39 % of the registered dairy cows, followed by Jersey, contributing about 11 % of the registered dairy cows, and then Ayrshire constituting less than 5 % of the registered dairy cows in South Africa (Waal & Blom, 2020).

So many industries suffered the effects of COVID 19, and the dairy industry was no exception. The 2021 first quarter (Q1) report of the Department of Agriculture, Forestry and Fisheries state that South Africa's milk production declined by 17 %, from 1044 billion litres to 866 million litres, in the 2021 Q1 compared to an 8 % increase to 1 044 billion litres reported in the 2020 fourth quarter (Phahlane et al., 2021). Compared to the previous quarter, the average producer price of milk per litre increased by 7.2 % in 2021 (Phahlane et al., 2021). The primary reason for the drop in production and the price increase was the sharp increase in price for some of the raw materials required, especially maize and soya beans (South African Milk Processors' Organization, 2021). Uncertainties regarding the effect of COVID-19 on the economy and dairy product demand also contributed to the drop in milk production and the increase in milk prices.

Regardless of the drop in milk production, the exportation of milk and cream increased by 16.7 % for the 2021 Q1. Similarly, the importation of the same product also increased by 29.7 % during the same period (Phahlane et al., 2021). It was, however, estimated that in 2021, South Africa was a net exporter of milk and cream, as well as buttermilk and yoghurt.

It is indisputable that COVID 19 resulted in job losses, including in the dairy sector. In most farms, this, together with the rise in feed costs, necessitated the culling of less productive cows to have a smaller herd of highly productive cows, which can also be easily managed. There were about 1.6 million dairy cattle in 2019 before COVID 19 (Visser et al., 2020); however, the herd size had dropped to about 1.2 million dairy cows after COVID 19 (Smith, 2021). There is an increased need for predictive tools to know which cows to cull and which ones to keep.

2.3. Milk yield prediction

Every dairy farmer desires to have foresight of how much milk each cow in the herd will produce. Accurate prediction of dairy cow milk yield is helpful to dairy farmers for various reasons. Milk yield prediction allows the farmer to decide which cows to keep for breeding purposes (Gorgulu, 2012; Getu & Misganaw, 2015). Those producing more milk can be kept for breeding as they are more likely to produce high-yielding progeny. Milk yield prediction also helps quickly identify genetically superior bulls based on their progeny's predicted milk yields (Gorgulu, 2012). This means that the collection of semen from these genetically superior bulls can commence sooner, and more inseminations can be done. Prediction information, together with milk estimated breeding values, enhances the breeding decision-making process (Getu & Misganaw, 2015).

Early detection of low-producing animals is also essential for culling decisions (Gorgulu, 2012). This is advantageous for economic reasons. It saves money that could have been wasted on rearing and maintaining an unproductive cow. Milk yield prediction is also essential for financial planning as it indicates how much feed the farmer will need and how much milk he expects from his herd (Gorgulu, 2012). Milk yield prediction helps allocate feed for an individual cow and a herd. Without the milk yield predictions, it is challenging to plan a physiologically balanced ratio that suits the needs of the animals (Jeretina et al., 2016).

Knowing the deviation of the observed milk yield values from the expected milk yield can be an indicator of mastitis and other diseases (Jensen et al., 2018). When a cow is mastitic, its milk production is most likely to drop. Therefore, an accurate prediction of milk yield gives an early insight into the cows that may be mastitic or poor milk producers. Milk yield prediction is, therefore, vital for better management as it enables better feeding, mating and culling decisions and disease detection (Jensen et al., 2018; Nguyen et al., 2020; Kliš et al., 2021).

2.3.1.Components used to predetermine milk yield

Driven by the farmers' desire to predetermine the milk yield (MY) each cow in the herd produces, various prediction components are used. These include genomic prediction, pedigree and progeny information, cow's own earlier milk yield records, and linear conformation traits correlated with milk yield. These are then implemented using statistical, technological or mathematical models such as linear regression to predict milk yield. Therefore, a prediction method is a combination of the prediction component or tools, e.g. use of genomics, and the prediction model, e.g. linear regression, being used. Prediction methods, therefore, come in different permutations. Also, with the advances in technology, there is room for more prediction methods. Both the tools and the models play an essential role in the accuracy of a prediction method, hence the need to try out various models on the aforementioned milk yield prediction tools. This section looks closely at the various prediction tools, their current ramifications, accuracies using different models and, lastly, possible gaps and permutations.

2.3.1.1. Genomic prediction

Genomic prediction is the use of single nucleotide polymorphisms (SNPs) throughout the whole genome instead of a quantitative trait locus (QTL) to explain variability in animals. Genomic prediction is complexly intertwined with other components such as pedigree and progeny information. This is because pedigree or progeny cattle can be genotyped to estimate breeding values. The genomic prediction uses two sources of information: genetic relationships among individuals and linkage disequilibrium between SNPs and QTL. The emphasis on these information sources varies with the genomic prediction statistical method used. Genomic best linear unbiased prediction (GBLUP) and Bayes are the two main methods used to carry out genomic prediction. The genomic best linear unbiased prediction exploits the relationship in a given population more comprehensively by quantifying the variation in associations between sibs and the historical connections between individuals in the base generation through the genomic relationship matrix. However, Bayes, a nonlinear method, can better exploit the linkage disequilibrium information gained through QTL mapping than GBLUP.

Genomic prediction is advantageous mainly because of its generally high prediction accuracy. Genomic prediction, especially though GBLUP can bring out an accurate estimate of the proportion of the genome shared by related individuals even when based on a smaller subset of markers. Therefore, it provides higher estimation accuracy for breeding values than estimates based on pedigree information alone (Habier et al., 2007)(Forni et al., 2011; Vallejo et al., 2017). The high accuracy results in shorter generation intervals through higher

contributions from young genetically superior bulls and heifers and increased selection intensity. These factors result in an increase in genetic gain for economically important traits. Genomic selection is pertinent when phenotypes are impossible, complicated or expensive to measure. Genomic selection also allows for early selection using estimated breeding values before the heifer has reached the age to produce milk, thus having economic benefits.

Genomic prediction accuracy is dependent on various other parameters. These include the training population size, range of genome linkage disequilibrium (LD), the relationship between training and validation data sets, the heritability of the trait, and genetic architecture of the trait, including the size of allele substitution effects at QTL (Daetwyler et al., 2008; Andonov et al., 2017). Also, new animals need to be regularly added to the training sample because the estimated associations between the SNP and the phenotype determining genes may be lost due to recombination or mutation. Another disadvantage is that even though the cost of genomic selection is decreasing, it is currently still high. Also, methods such as Bayes in the genomic prediction method are computationally demanding. Genomic prediction also requires expert knowledge and personnel, which is often not readily available on the farm; hence, farmers use other simpler methods and only make genomic predictions for high-profile cattle.

Accuracies and error measurements for genomic prediction include estimated breeding values (EBV). Prediction is, therefore, dependent on the heritability of the trait, making genomic prediction best suited for breeding purposes since it is rudimentarily based on genetic factors and not environmental factors (both temporary and permanent environment). Without estimating breeding values, it is not easy to compare accuracies of genomic prediction with other prediction methods that fully account for both genetic and environmental effects on a particular trait for an individual. Although it may seem low, the accuracy of genomic prediction is more reliable from a breeding perspective since it is solely based on genes. Some analyses used genomic prediction to predict milk yield.

Ding (2013) measured accuracy as the correlation between two EBVs, i.e. GEBV and the conventional EBV of the progeny-tested bulls. The GBLUP had the highest accuracy of 0.76, followed by the conventional BLUP with 0.75, and BayesB had an accuracy of 0.73. Zhang et al. (2022) compared the pedigree-based best linear unbiased prediction (PBLUP) and the single-step genomic best linear unbiased prediction (ssGBLUP) for milk yield. The reliability of animal breeding values was calculated as the R^2 value between genomic estimated breeding and actual values. The R^2 value for the GEBV calculated using ssGBLUP was higher than that of EBV calculated through the PBLUP in the genotyped and whole populations. Using the multiple trait model for the genotyped subpopulation, the R^2 was 0.81 for the PBLUP and 0.83

for the ssGBLUP. For the multiple trait model on the whole population, the R^2 value was 0.61 and 0.62 for the PBLUP and the ssGBLUP, respectively. (Zhang et al., 2022)

2.3.1.2. Earlier milk yield records

Another prediction method is using earlier milk yield records to predict the milk yield for the whole lactation. Earlier milk yield records together with general trends of milk yield based on breed, region, season, parity and lactation number from Automatic Milking Systems (Kliś et al., 2021). The methods of implementing or analyzing the records may differ. Some may use artificial neural networks (ANN) and mathematical formulas based on variables such as season.

Gorgul (2012) used the earlier milk yield records through ANN to predict milk yield in Brown Swiss cattle. The results from the ANN were then compared with the multiple regression statistical method. For the ANN, the best results were found from the first four test-day records, with the prediction from the fourth day having an R^2 value of 0.90 and a root mean square error (RMSE) of 454.4. Multiple linear regression had a much lower R^2 value of 0.47 and an RMSE of 1247.9 (Gorgulu, 2012). This research proved the possibility of using earlier milk yield records to predict the 305-day yield. The use of ANN in milk yield prediction also proved viable.

Jingar et al. (2014) created 305-day milk yield trends for normal cows and cows with mastitis. The models were then used to predict milk yield for normal or mastitic cows. The statistical analyses were done using the Gamma type function. For the first parity cows, R^2 values were 0.85 for normal cows and 0.84 for mastitic cows (Jingar et al., 2014).

Liseune et al. (2020) used milk yield records to interpolate, backtrack, find values on a missing window or predict future milk yield through deep learning; precisely, a sequential autoencoder (SAE). The model used milk yield patterns for the whole lactation to estimate latent values. The SAE was then compared with the conventional multilayer perceptron model (MLP), which inputs herd and parity information and lagged milk yields. The 305-day prediction performance test results widely varied based on the number of days with missing records. With only results from 30 days, the highest RMSE for milk yield was 1231.5 kg, and the lowest R^2 value was 0.64. With results from 270 days, the least RMSE was 79.43, and the highest R^2 value was 1.00 (Liseune et al., 2020).

2.3.1.3. Linear-type traits

Another tool used in the prediction of milk yield is linear-type traits. Harris (2015) defined linear-type traits as objective descriptions of an animal's body. These describe the degree of

each trait, not the desirability (ICAR, 2015). Linear-type traits are lowly to moderately heritable and are linked with functionality (Getu & Misganaw, 2015). Hence their use in making predictions for quantitative traits. Since linear-type traits are visible and easier to measure, they can be important indicators of production traits with which they are correlated.

The International Committee for Animal Recording (ICAR) recognizes stature, chest width, body depth, angularity, rump angle, rump width, rear legs rear-view, rear legs side-view, foot angle, fore udder attachment, front teat placement, teat length, udder depth, rear udder height, central ligament (median ligament) and rear teat position. These linear-type traits can be classified into four groups, i.e. udder conformation, feet and leg conformation, thoracic and abdominal body conformation, and rump and loin structure traits (Atkins et al., 2008; Getu & Misganaw, 2015).

2.3.1.3(a). Udder conformation traits

Udder conformation traits include fore udder attachment, front teat placement, teat length, udder depth, rear udder height, central ligament (median ligament) and rear teat position. Udder conformation includes a detailed description of the udder's suspensor apparatus (Atkins et al., 2008).

The rear udder width is measured at the crease formed when the udder meets the legs. Rear udder width indicates udder capacity. Hence, the wider the udder, the higher the capacity (Stamschror & Card, 2000). Figure 2.1 shows various udder widths.

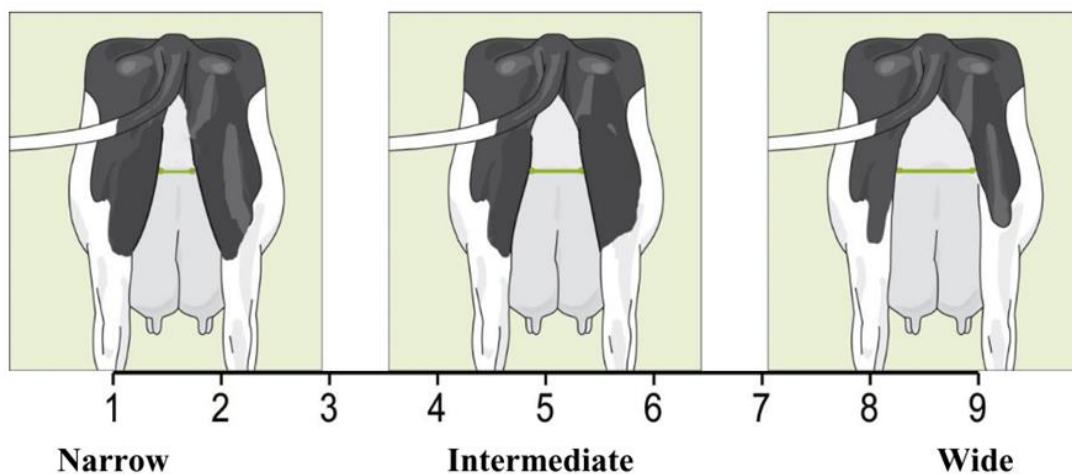


Figure 2.1: Cow images showing various rear udder widths. Source: ICAR (2015).

Udder depth measures the udder floor depth relative to the hock. Udder depth increases with age (Stamschror & Card, 2000). Higher udders are related to fewer incidences of mastitis and fewer udder injuries. Figure 2.2 shows how udder depth is viewed.

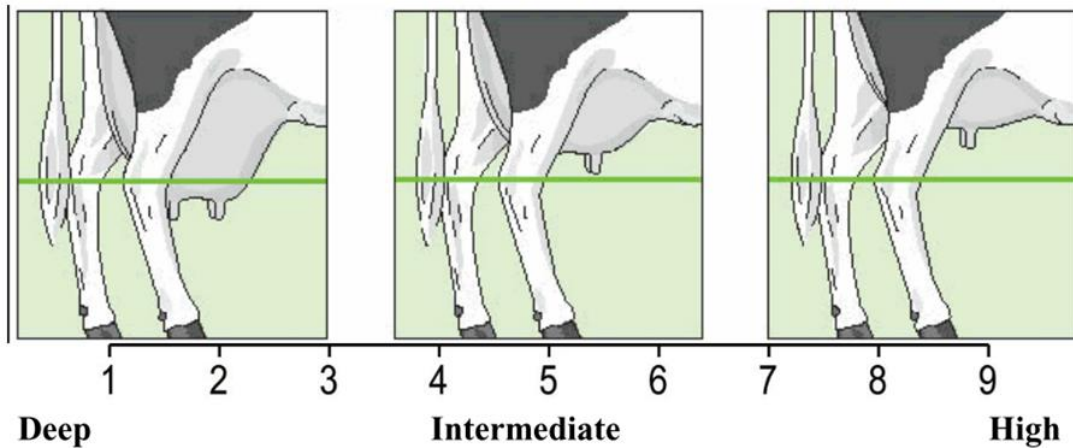


Figure 2.2: Cow images showing various udder depths. Source: ICAR (2015).

Rear udder height is the distance between the bottom of the vulva and the top of the milk-secreting tissue in relation to the animal's height. A high rear udder attachment indicates more udder capacity (Stamschror & Card, 2000). Figure 2.3 shows cows with various rear udder heights.

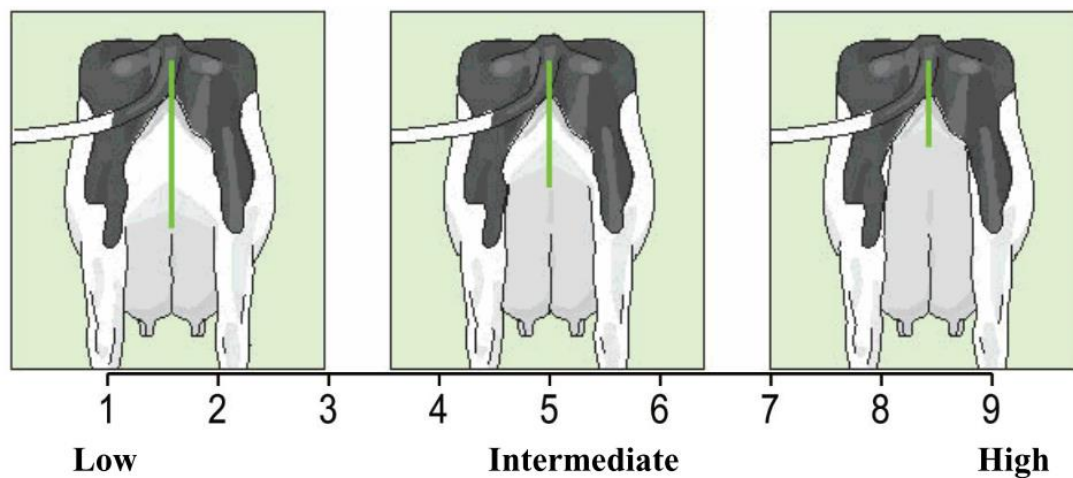


Figure 2.3: Cow images showing various rear udder heights. Source ICAR (2015).

Fore udder attachment indicates the fore udder attachment strength to the abdominal wall by the lateral ligaments. A moderate bulge of the fore udder is associated with high-producing dairy cows (Stamschror & Card, 2000). The udder should have minimum pendulation when the cow is walking. Figure 2.4 shows variations in fore udder attachment.

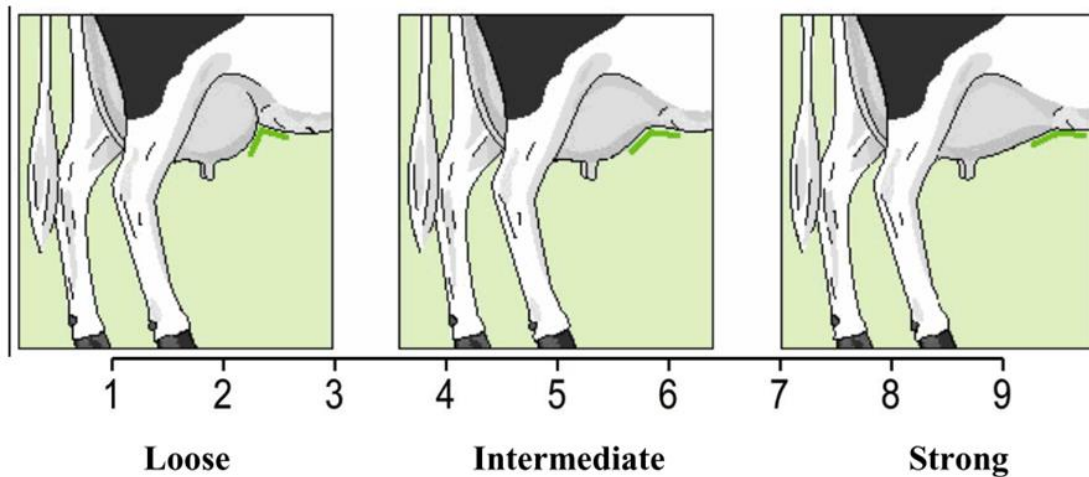


Figure 2.4: Cow images showing various fore udder attachments. Source: ICAR (2015).

2.3.1.3(b). Thoracic and abdominal conformation traits

Thoracic and abdominal confirmation traits include chest width, angularity, body depth and stature. Thoracic and abdominal capacity, along with angularity, facilitates the ability of dairy cows to process large quantities of roughage and maintain high milk production (Getu & Misganaw, 2015)

Angularity refers to the ribs' angle and spring/degree of openness. The cow is angular when the ribs point towards the udder. If the ribs are difficult to see and the cow has no rib opening, the cow is said to be coarse (ICAR, 2015). Figure 2.5 shows cows with various angularities.

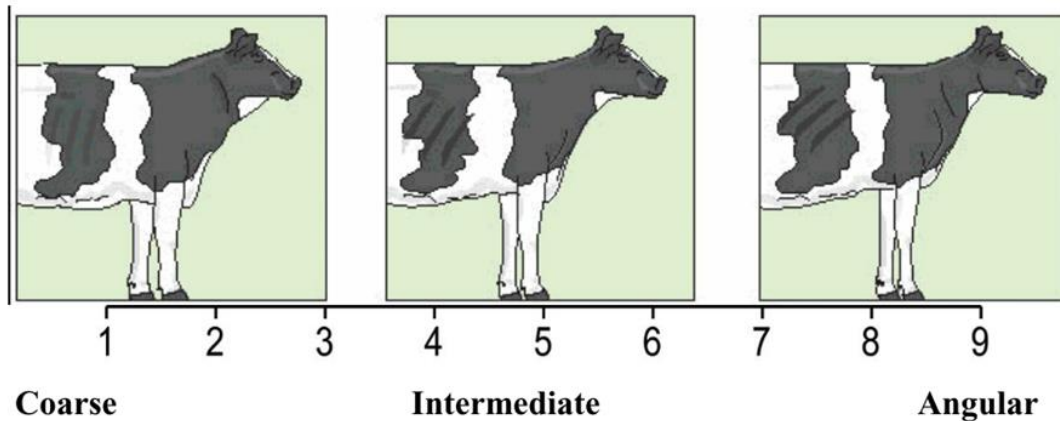


Figure 2.5: Cow images showing various angularities. Source: ICAR (2015).

Body depth is the distance between the top of the spine and the bottom of the body at the beginning of the last rib. Stamschror & Card (2000) claimed that there is no evidence for the association between body depth and the capacity to consume large amounts of forage and produce more milk (Stamschror & Card, 2000). More recent research, however, shows a correlation between body depth and milk yield, albeit low (Alphonsus et al., 2010; Tapki, 2013). Campos et al. (2015) stated that body depth is more linked to milk fat than milk yield. Figure 2.6 shows cows with different body depths.

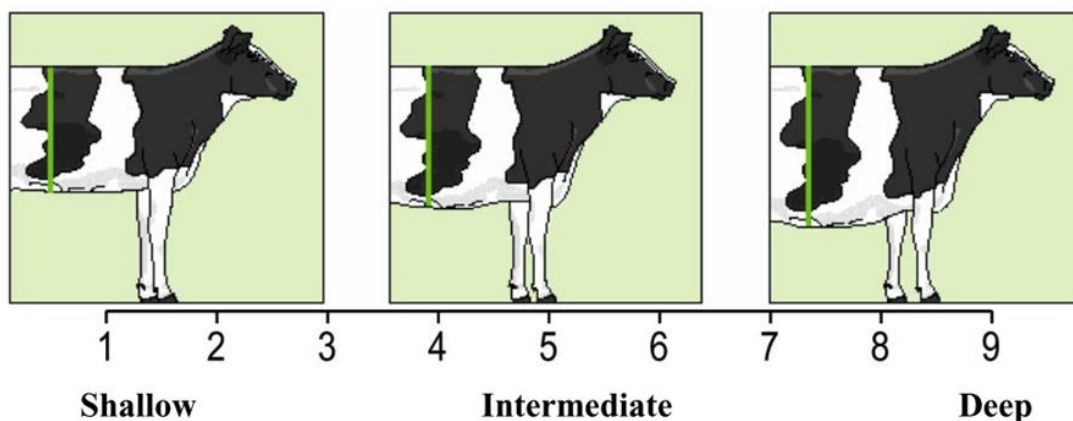


Figure 2.6: Cow images showing various body depths. Source: ICAR (2015).

Stature is measured from the top of the spine to the ground, just between the hips (ICAR, 2015). Stature is positively correlated with milk yield, i.e. the taller the animal, the higher the milk yield (Gocheva-ilieva & Yordanova, 2022). Figure 2.7 shows how stature is measured.

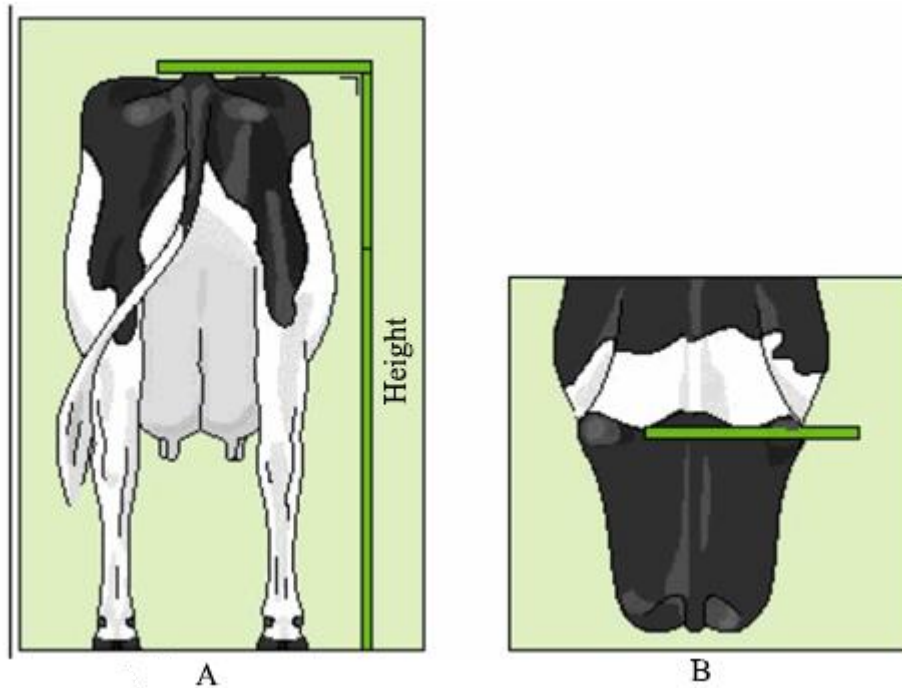


Figure 2.7: Cow images showing how stature is measured. Image A shows the rear view of how stature is measured and Image B shows the point of measurement. Source: ICAR (2015).

2.3.1.3(c). Rump and loin structure

The rump and loin structure fastens the cow's abdominal and lumbar regions to her mammary system, feet and legs. With frailty in this area, a cow's productive life is thwarted (Atkins et al., 2008). Two major linear-type traits fall into this category, rump width and rump angle. Rump angle is more important as an indicator of reproductive performance and health. Cows with higher pin bones are more vulnerable to dystocia and are more prone to vaginal infections because the vagina cannot drain effectively (Getu & Misganaw, 2015). Rump width is important in milk yield prediction. Rump width is the distance between the most posterior point of pin bones. Figure 2.8 shows three cows with different rump widths.

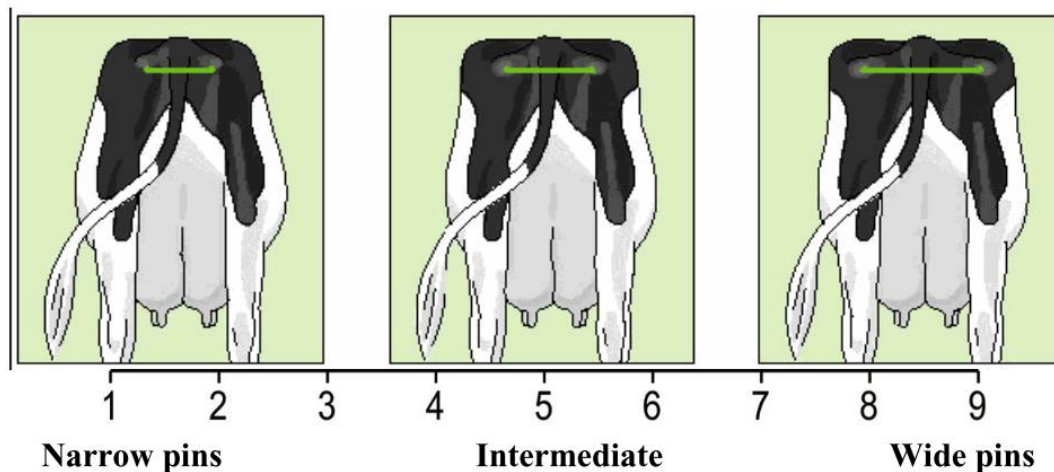


Figure 2.8: : Cow images showing various rump width. Source: ICAR (2015).

2.3.2. Correlation between linear-type traits and milk yield

The basis for prediction is the relationship between these linear-type traits and milk yield. Therefore, it is salient to understand the correlations between conformation traits and milk yield. This aids in understanding which conformation traits should be given priority when predicting milk yield.

Linear-type traits are not only linked with milk yield but with a vast number of quantitative traits. For example, Wall et al. (2005) found a positive correlation between legs, feet and fertility. Parke et al. (1999) found positive correlations between body weight and body capacity, stature and rump width.

One of the traits widely accepted to correlate with milk yield is angularity. The angularity and milk yield phenotypic correlations range from 0.10 to 0.31 (Tapki, 2013; Bohlouli et al., 2015; Campos et al., 2015; Harris, 2015; Janković et al., 2020). The more angular a cow is, the higher the milk yield. Kern et al. (2015), however, found no correlation between angularity and milk yield. Khan & Khan (2016) also found no correlation between MY and angularity. The discrepancies in the findings could be attributed to possible variations in visual judgments of angularity.

Another conformation trait of significance when considering milk yield is udder depth. Udder depth is negatively correlated with milk yield (Campos et al., 2015; Khan & Khan, 2016; Janković et al., 2020). An udder lower than the hocks exposes the udder to damage and

infection, resulting in reduced longevity, reduced milk yield and increased susceptibility to mastitis (Stamschror & Card, 2000).

Unlike udder depth, fore udder attachment, rear udder height, rear udder width, central ligament and teat length are positively correlated with milk yield (Campos et al., 2015; Khan & Khan, 2016; Janković et al., 2020). A wide udder with a high milk-secreting tissue signifies a higher capacity for milk, hence increased milk yield. Udder support is vital as it bears on udder strength and teat placement. For example, if fore udder attachment to the abdominal wall is weak, the udder may become too pendulous when walking, increasing damage. A broken central ligament causes the teats to face outside. Therefore, the hind legs may damage the teats when the cow walks. Therefore, all udder attachments to the ventral abdominal wall and the pelvic floor are pivotal to udder health and, eventually, milk yield (Atkins et al., 2008). Tapki (2013) reported a correlation of -0.23 between fore udder attachment and MY. Bohlouli et al. (2015), on the other hand, found no correlation between RUW and MY.

Body depth is another essential indicator of how much milk yield a farmer expects from a particular cow. The deeper the body, the higher the animal's capacity. In most literature, the correlation between body depth and milk yield ranged from 0.10 to 0.27 (Alphonsus et al., 2010; Tapki, 2013). Other traits of importance are rump width, with a moderate correlation of about 0.25 (Harris, 2015) and stature, ranging from 0.17 to 0.65 (Alphonsus et al., 2010; Harris, 2015; Khan & Khan, 2016). There is a weak correlation (<0.2) between milk yield and chest width or with foot angle (Alphonsus et al., 2010; Tapki, 2013; Bohlouli et al., 2015; Campos et al., 2015; Harris, 2015; Khan & Khan, 2016).

2.3.3. Milk yield prediction using linear-type traits

Based on significant correlations, Yakubu (2011) tested heart girth, fore right teat length, fore left teat length, rear right teat length, rear left teat length, udder circumference, and udder height to predict milk yield. The R^2 values of the individual traits in predicting milk yield were inferior ($R^2 < 0.2$). Predictions from all the traits combined gave a moderate R^2 value of 0.69 (Yakubu, 2011).

Gocheva-ilieva & Yordanova (2022) examined the predictability of milk yield from stature, chest width, rump width, rear legs (rear-view), rear leg set (side-view), hock development, bone structure, foot angle, foot depth and udder width. The list, however, ignored angularity and rear udder height (Campos et al., 2015). Locomotion, farm and lameness, which are not conformation traits, were included. Among the linear-type traits examined, Gocheva-ilieva

and Yordanova (2022) found udder width, chest width, and stature to be the three essential conformation traits in predicting milk yield. When all traits were considered, the R^2 values ranged from 0.94 to 0.95.

Ozkaya (2015) showed huge opportunities in using image data to make predictions. Using udder depth, udder width and udder area, the R^2 value was highest when all the trait measurements were used: $R^2= 0.66$. The highest R^2 value from a single trait was found for the udder area ($R^2=0.61$).

Ozkaya's method, like any other prediction method based on linear-type traits, required the establishment of correlations between the conformation traits and milk yield as this guided how much weight he put on a particular trait. This is, however, a challenge because the extent of the correlations between a particular trait and milk yield differs widely.

With deep learning, images do not need to be taken at the same distance from the cow. Taking images from the same distance may include the distance from the desired object as a variable/parameter that affects prediction. What is ideal is to train a deep learning model with images having different distances from the cows. This trains the model to estimate trait measurements based on the features around it. With deep learning, less precision is required in taking the pictures.

2.4. Computer vision deep learning

Computer vision deep learning is a combination of the concept of computer vision (CV) and deep learning (DL). Both concepts fall under the field of artificial intelligence (AI). Brownlee (2019) defined computer vision as a field of study seeking to develop techniques to help computers "see" and comprehend the content in visual data such as photographs and videos. On the other hand, deep learning is a type of machine learning, inspired by how the human brain works, based on artificial neural networks with feature/representation learning. Machine learning is a subset of AI involving the study of computer algorithms that can automatically improve using data or experience (O'Mahony et al., 2020). Representation learning is a system that allows automatic location and learning of the features needed for object classification or detection from raw data, allowing a machine to both learn the features and use them to perform specific tasks. Artificial neural networks, also called neural networks are intelligent systems that can solve various tasks, including pattern recognition, optimization and prediction. Figure 2.9 shows a schematic relationship between computer vision, deep learning, machine learning and AI.

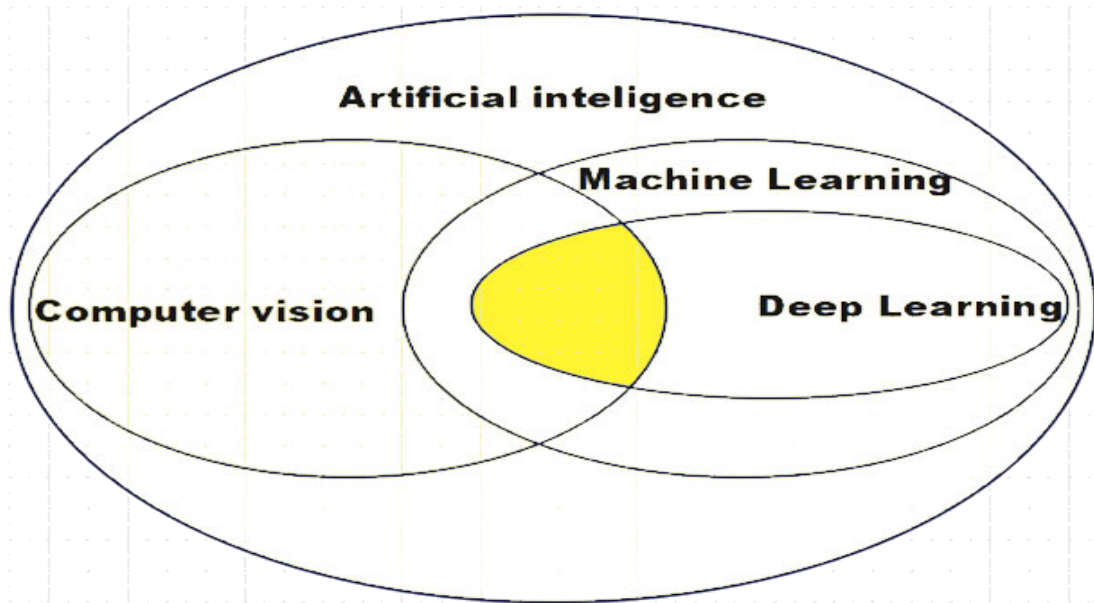


Figure 2.9: Relationship between computer vision, deep learning, machine learning and Artificial Intelligence. (Johnson 2019)

Deep learning enables computer vision systems to attain greater accuracy in tasks such as image classification, object detection, semantic segmentation and simultaneous localization and mapping (SLAM) (O'Mahony et al., 2020). Applications of DL require minimum expert analysis and fine-tuning since these deep learning networks train themselves by finding patterns in the data given to them and their associated labels.

2.4.1. Convolutional neural networks

Several neural networks are used in deep learning. Convolutional neural networks (CNN) are, however, the best for computer vision as they are specially built to work with images. Convolutional neural networks can be divided into two stages; feature extraction and classification of the image. Figure 2.10 is a schematic representation of the flow of convolutional neural networks.

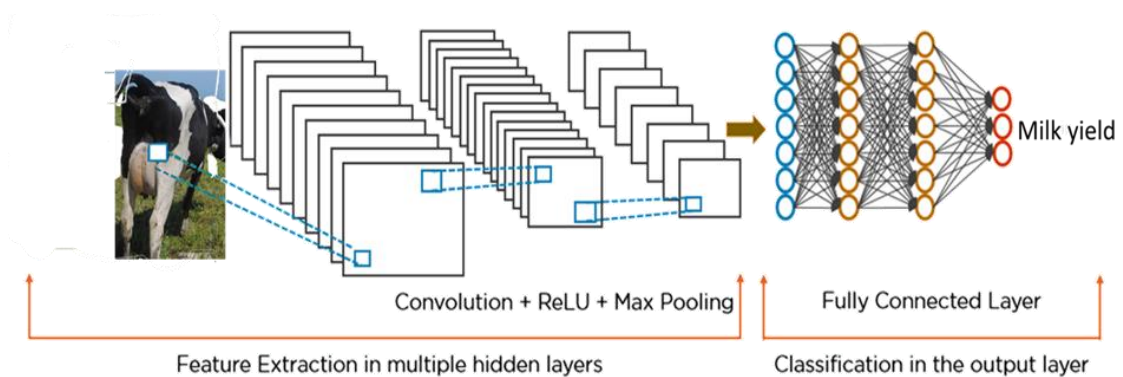


Figure 2.10: A schematic representation of the building blocks of a convolutional neural network.

Deep learning supports flexibility by re-training pre-trained CNN frameworks using a custom dataset for any use case (O'Mahony et al., 2020). This process is called transfer learning. A pre-trained model is a saved network that, previously, was trained on a large-scale image-classification task with a large data set.

There are various families of pre-trained models, namely Visual Geometry Group (VGG), Inception (GoogLeNet), ResNet and EfficientNet (Hassaballah & Awad, 2020). For this research, a modification of the Inception architecture was used, called the Xception network (Chollet, 2017). Francois Chollet defined it as a deep convolutional neural network architecture based on depthwise separable convolutions (Chollet, 2017).

2.4.2. Challenges of computer vision deep learning

The accuracy of deep learning comes with a cost of an increased requirement for data and processing power. It is vital to have a dedicated hardware such as a high-powered graphics processing units or GPUs (O'Mahony et al., 2020).

Deep Learning requires vast amounts of data. For example, the ImageNet consists of 1.5 million images with 1000 object categories, the Pattern analysis, statistical modelling, and computational learning visual object classes (PASCAL VOC) Dataset consists of 500K images with 20 object categories and Microsoft Common Objects in Context (COCO) consists of 2.5 million images with 91 object categories (Arai et al., 2014).

Due to the vast amount of data required, training a DNN takes a lengthy process. Depending on computing hardware availability, training can take hours or days. Moreover, training for any given application often requires many iterations as it entails trial and error with different training parameters. The most common technique to reduce training time is transfer learning. (O'Mahony et al., 2020)

Vision processing results using DL are also reliant on image resolution. For example, obtaining satisfactory performance in object classification also requires high-resolution images or videos. The high resolution also taxes on the consequent increase in the amount of processed, stored, and transferred data. Image resolution is significantly pertinent in applications that necessitate object detection and classification at a distance. The frame reduction techniques such as using the Scale-Invariant Feature Transformation (SIFT) features

first to identify the region of interest are helpful for image resolution and reduce the time and data required for training.

2.4.3. Uses of computer vision in animal production

Various applications require visual classification or judgement in the field of animal science. These include judgements on body condition scoring, animal behaviour, carcass fat deposition, meat marbling, or classification of eggshell quality. Aided by lenses, such as microscopes, it is possible to evaluate cell morphology in a blood smear or spermatozoid motility and defects. Additionally, other signals such as X-rays, ultrasounds and infrared are extensively used to produce images for diagnostic purposes. Most of the conventional methods used for the optical measurement of traits of interest require expert personnel to be trained from time to time to maintain good measurement quality (Fernandes et al., 2020).

Furthermore, most of these measuring processes are time-demanding, pricey for the farmer and stressful to the animals. Therefore, there is an ongoing interest in developing automatic and indirect methods of measuring traits of interest and monitoring livestock. For such tasks, computer vision deep learning generally uses algorithms and principles of pattern recognition, image analysis, and processing to tackle the most diverse problems. The good thing about CV DL is that it constantly adjusts its perception to perform better as it captures the information. (Fernandes et al., 2020).

Automated CV systems enable extensive phenotyping in livestock, and the data created by such systems can be used for many different applications, from developing intelligent farm management tools to advancing breeding programs (Fernandes et al., 2020). Some of the CV systems that are already being implemented in animal science will be discussed in this section.

2.4.3.1. Evaluation of body condition, carcass and meat traits

Some of the first applications of CV systems were in meat sciences, with the earliest reported studies found in the 1980s (Cross et al., 1983; Wassenberg et al., 1986). In these studies, the following components constituted the system, a camera, a light source, a digitizer, and a computer. An operator was required to position the meat cuts at a known angle and distance from the camera, with the same background and illumination, then capture the images. Then, the interest was to predict the cut content of lean meat and fat and compare the results from the CV systems to trained meat graders. The prediction equations developed with the system-measured variables, albeit not fully automated, presented slightly better results (R^2 from 0.93-

0.95) than the prediction equations developed with variables measured by trained graders (R^2 from 0.84-0.94) (Cross et al., 1983; Wassenberg et al., 1986).

Since these earlier studies, computer vision has been on the rise in predicting various meat quality traits, not only for beef but also for poultry, fish, and pork. Modern applications are more automated, using computer vision machine learning techniques to determine the meat's crude protein, fat content, and more advanced chemical traits such as fatty acids profile and freshness (Wang et al., 2013; Khoshnoudi-Nia & Moosavi-Nasab, 2019). Moreover, applications include the prediction of meat quality, tenderness, palatability, and other traits customarily evaluated by a panel of trained personnel (Luis Nunes et al., 2015; Luis Nunes et al., 2015; Zapotoczny et al., 2016) or even automatic sorting and weighing cuts and viscera which is typically performed manually (Paluchowski et al., 2016; Adamczak et al., 2018).

The evaluation of meat and carcass traits in live animals is not an easy task to perform. Doeschl et al. (2004) attempted to predict carcass composition in live pigs' foreloin and hind loin regions. A computer vision system attained a predictive R^2 of 0.31 and 0.19 for fat and 0.04 and 0.18 for lean meat from the foreloin and hind loin, respectively. These were undisputably low R^2 values, indicating poor predictability of fat or lean meat from live pigs. In a quest for performance improvement in predicting meat and carcass traits on live animals, researchers shifted to using medical imaging devices such as ultrasound and computed tomography (CT) scanners for such tasks. Carcass measurements using ultrasound improved the correlations of 0.6 and 0.56 for lean meat and fat depths, while carcass measurements using CT showed correlations between 0.48–0.67 for fat and 0.91–0.94 for lean meat (Doeschl et al., 2004; Lucas et al., 2017).

These medical devices have their disadvantages concerning animal handling and cost. Therefore, some recent works developed CV systems based on 3D cameras for automated non-contact estimations of muscle score (Alsahaf et al., 2019) and of fat and lean muscle content (Fernandes et al., 2020) on live pigs. Alsahaf et al. (2019) developed a system that could predict muscle scores between 1 and 5 using morphometric features extracted from the images of moving pigs. With a gradient boosted classifier, a classification accuracy between 0.3 and 0.58 and a mean absolute error of 0.65 was achieved. Meanwhile, Fernandes et al. (2020) used computer vision deep learning methods that do not require image processing. The deep learning approaches achieved better results, R^2 of 0.50 for lean muscle depth and 0.45 for fat depth, compared to previous studies that did not use deep learning. Nevertheless, these R^2 values are still low, indicating poor prediction accuracy. Better results on lean muscle and fat predictions are achievable (Fernandes et al., 2020).

2.4.3.2. Animal tracking and behaviour analyses

Computer vision systems are also used for animal tracking, monitoring, and identification of changes in their daily behaviour. Animals tend to synchronize their behaviour within a group, making it possible to notice apparent deviations resulting from environmental stress, management problems, or disease. Therefore, researchers desire to understand behavioural changes in animals and their relationship with other traits of interest, such as animal health status and growth. Conventionally, trained evaluators conduct closer evaluations of animal behaviour and health during vaccination times or the transfer of animals from one location to another. This is because managers and workers often have limited time to observe a group of animals.

Computer vision systems are used to acquire and store images and videos that the farmers can assess later or remotely. This eases the burden on management since the evaluator does not need to be physically present, which otherwise can cause behavioural changes in the animals. Some computer vision systems can automatically classify animal behaviour and alert the manager in real-time regarding essential changes (Barnard et al., 2016; Matthews et al., 2017; Psota et al., 2019). Tracking can be done for a wide range of animals, such as poultry (Sassi et al., 2016; Li et al., 2020), fish (Saberioon et al., 2016), cattle and pigs (Nasirahmadi et al., 2017).

To identify behaviours such as drinking and feeding, manual segmentation of the regions of interest (ROI) from the captured images proves to be effective (Kashiha et al., 2013). The aim is to identify the animals and other objects, such as the water source and feeders, so it is possible to track the interaction between the animals and these objects (Fernandes et al., 2020).

2.4.3.3. Identification of mastitis and digital dermatitis by thermography

Most computer vision applications use a standard charged-coupled device or an active-pixel sensor cameras. However, various other sensors are of interest, such as thermal imaging and depth cameras. Thermal cameras are typically used in veterinary sciences for diagnosis in clinical examinations. Thermal images are used to identify differences in skin temperature, which can be related to inflammatory processes, necrosis, infection, stress, and overall health (Fernandes et al., 2020). Infrared thermography (IRT) has been used in research to identify mastitis in dairy cattle and sheep (Hovinen et al., 2008; Martins et al., 2013), for detecting

udder temperature increase after inoculation with *E. coli* (Metzner et al., 2015) and for digital dermatitis in sheep (Byrne et al., 2018).

Animal skin temperature can differ based on external factors, such as the environmental temperature, wind speed, camera positioning, and body region (Fernandes et al., 2020). Thus, IRT applications can not fully be automated. More research should be done on developing automated IRT methods of measurements under farm conditions (Fernandes et al., 2020).

2.4.3.4. Prediction of individual body measurements and body condition scores using three-dimensional cameras

Three-dimensional (3D) cameras can measure traits in the 3-dimensional space, such as an animal's body position, gait, and volume. Within 3D images, it is easier to estimate measurement traits based on the camera's distance to the animal (Fernandes et al., 2020). It is possible to measure traits such as a cow's body condition score (BCS) using 3D cameras. Spoliansky et al. (2016) developed an automatic dairy cows' BCS system using 3D cameras (Spoliansky et al. 2016). This research captured top view images from cows leaving the milking parlour. The images were then automatically processed through background removal, cow centralization on the image, and normalization. Feature extraction was done from the processed images, and the extracted features were used to develop multiple linear regression models via stepwise regression. Even though correlations between the variables extracted and BCS were low, the developed model had comparable or better R^2 values (0.68) to previous studies based on manual image processing using either standard digital images (Bewley et al., 2008; Bercovich et al., 2013) or thermal cameras (Halachmi et al., 2013). Other applications in which 3D cameras show promising results are in estimating animal body measurements such as heights, widths, volume and area, as well as; body weight (Kongsro, 2014; Pezzuolo et al., 2018).

Fernandes et al. (2019) developed an automated real-time video processing and prediction CV system for body weight in live pigs using 3D cameras (Fernandes et al., 2019). The videos were collected under farm conditions. Using multiple linear regression models, they achieved high prediction accuracy ($R^2=0.92$) with features extracted from the images as predictor variables. An adaptation of Fernandes et al. (2019) 's CV system was also evaluated to predict body weight in beef cattle achieving an R^2 of 0.79-0.91 with an artificial neural network approach (Cominotte et al., 2020).

2.4.3.5. Prediction of milk yield from linear-type traits

Computer vision can also be used to predict milk yield based on linear-type traits. This is made possible because of the relationship existing between some linear-type traits and milk yield. Conformation traits are visible in an image, hence, the same judgement done from live visualisation can be done using an image or video. An image can be understood by computers, thereby making possible the automation of milk yield prediction. Ozkaya (2015) developed a system of milk yield prediction using cow udder images. When all traits were considered the R^2 value was 0.66. The uppermost R^2 value from a single trait was found for the udder area ($R^2=0.61$) (Ozkaya, 2015). The application of computer vision to make milk yield predictions is low and there is a need for more research in this area.

2.4.4. Advantages of using computer vision deep learning in predicting milk yield

Apart from the general advantages of milk yield prediction discussed in section 2.3, the use of image data to predict milk yield through deep learning has various advantages. Milk yield prediction can be done early in a cow's life when the cattle is still a heifer and milk production hasn't commenced. This is because the method is based on linear conformation traits visible even before lactation. This in turn enhanced the decision-making process in terms of mating, culling, feeding and general animal management (Gorgulu, 2012; Getu & Misganaw, 2015).

Prediction of milk yield using images is convenient to implement since images can be taken using cell phones which are currently ubiquitous. Also, once the model is trained and functional, prediction is easy since the process is automated. No expert knowledge is required to make the prediction.

2.5. Data augmentation and the augmentation types

The most common method to overcome the challenge of inadequate datasets and reduce overfitting deep learning models for image classification is to increase the dataset using label-preserving transformations. This process is known as data augmentation. Data augmentation involves the artificial generation of extra training data from the available ones, such as cropping, scaling, or rotating images (O'Mahony et al., 2020). The common techniques to generate new images are horizontal or vertical flips, stretching, rotation at some degrees, outward or inward scaling, shearing, padding, cropping and adding Gaussian noises. Figure 2.11a to 2.11i show illustrations of these augmentation types.

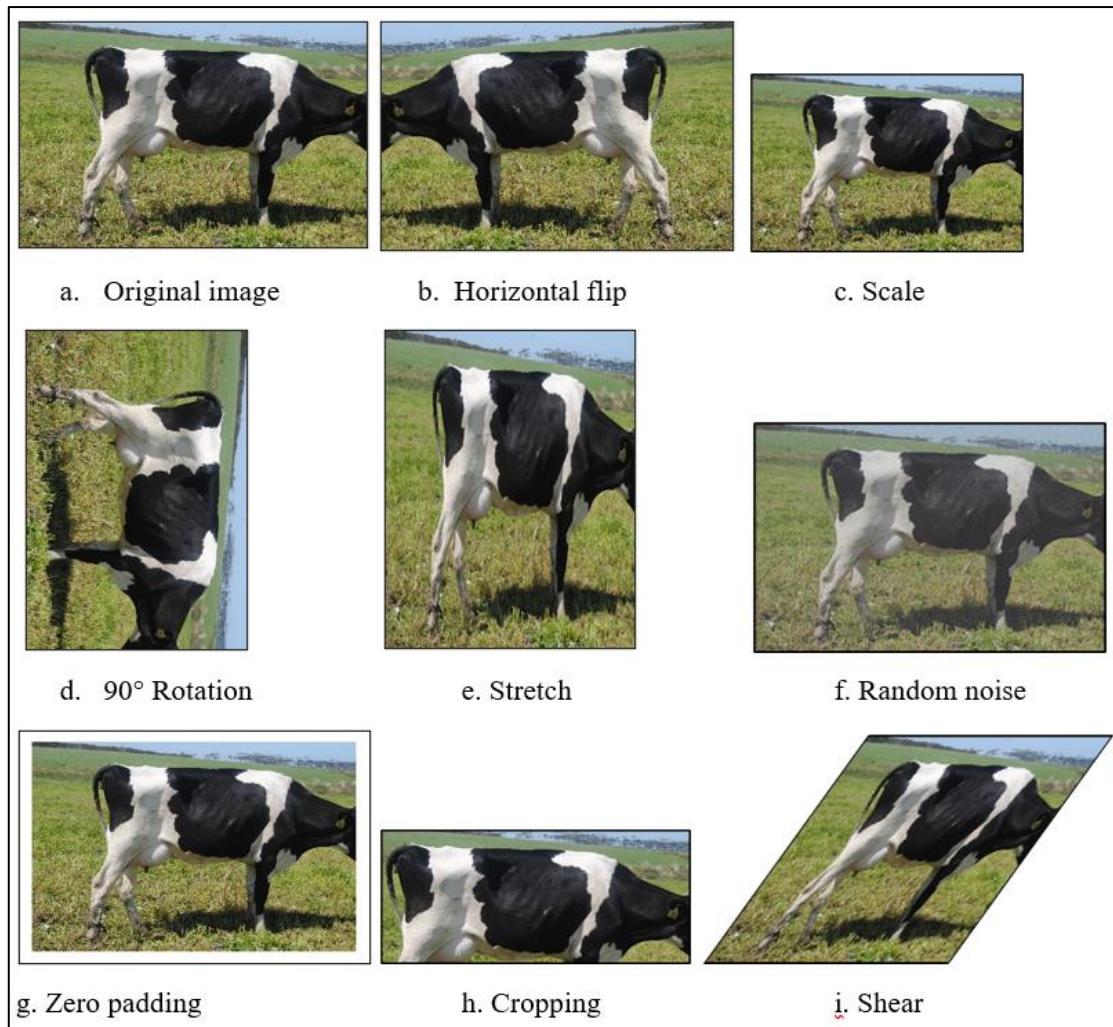


Figure 2.11: Representative images with data augmentation techniques, such as flipping (b), scaling(c), rotating (d), stretching (e), noise (f), padding(g), cropping (h), and shearing (i).

2.6. Summary

Prediction of milk yield is salient in the dairy industry as it enables better decisions regarding feeding, culling, mating and disease detection. There are various ways in which milk yield can be predicted, namely through genomics, pedigree and progeny information, cow's own earlier MY records, and linear conformation traits. These methods are implemented through various statistical tools, and one of these tools is deep learning. The broad objective of this research was to determine the predictability of milk yield using cows' visual images.

2.7. References

- Adamczak, L., Chmiel, M., Florowski, T., Pietrzak, D., Witkowski, M., & Barczak, T. (2018). The use of 3D scanning to determine the weight of the chicken breast. *Computers and Electronics in Agriculture*, 155, 394–399. Doi: 10.1016/J.COMPAG.2018.10.039
- Alphonsus, C., Akpa, G. N., Oni, O. O., Rekwot, P. I., Barje, P. P., & Yashim, S. M. (2010). Relationship of linear conformation traits with bodyweight, body condition score and milk yield in Friesian x Bunaji cows. *Journal of Applied Animal Research*, 38(1), 97–100. Doi:10.1080/09712119.2010.9707164
- Alsahaf, A., Azzopardi, G., Ducro, B., Hanenberg, E., Veerkamp, R. F., & Petkov, N. (2019). Estimation of Muscle Scores of Live Pigs Using a Kinect Camera. *IEEE Access*, 7, 52238–52245. Doi:10.1109/ACCESS.2019.2910986
- Andonov, S., Lourenco, D. A. L., Fragomeni, B. O., Masuda, Y., Pocrnic, I., Tsuruta, S., & Misztal, I. (2017). Accuracy of breeding values in small genotyped populations using different sources of external information—A simulation study. *Journal of Dairy Science*, 100(1), 395–401. Doi: 10.3168/jds.2016-11335
- Arai, K., Kapoor, S., & Conference, C. V. (2014). *Advances in Computer Vision*. In *Advances in Computer Vision* (Vol. 1, Issue Cvc). Doi: 10.4324/9781315802145
- Atkins, G., Shannon, J., & Muir, B. (2008). Using Conformational Anatomy to Identify Functionality & Economics of Dairy Cows. *WCDS Advances in Dairy Technology*, 20, 279–295.
- Barnard, S., Calderara, S., Pistocchi, S., Cucchiara, R., Podaliri-Vulpiani, M., Messori, S., & Ferri, N. (2016). Quick, accurate, smart: 3D computer vision technology helps assessing confined animals' behaviour. *PLoS ONE*, 11(7). Doi: 10.1371/JOURNAL.PONE.0158748
- 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. Doi: 10.3168/JDS.2013-6568
- Bewley, J. M., Peacock, A. M., Lewis, O., Boyce, R. E., Roberts, D. J., Coffey, M. P., Kenyon, S. J., & Schutz, M. M. (2008). Potential for Estimation of Body Condition Scores in Dairy Cattle from Digital Images. *Journal of Dairy Science*, 91(9), 3439–3453. Doi:10.3168/JDS.2007-0836
- Bohlouli, M., Alijani, S., Varposhti, M. R., Researchers, Y., Club, E., Branch, T., Delta, A.,

- & Company, G. (2015). Genetic relationships among linear type traits and milk production traits of Holstein dairy cattle. 15(4), 903–917. [doi:10.1515/aoas-2015-0053](https://doi.org/10.1515/aoas-2015-0053)
- Brownlee, J. (2019). Deep learning for computer vision: image classification, object detection, and face recognition in python. Machine Learning Mastery.
- Byrne, D. T., Berry, D. P., Esmonde, H., MCGovern, F., Creighton, P., & Mchugh, N. (2018). Infrared thermography as a tool to detect hoof lesions in sheep. *Translational Animal Science*, 3(1), 577–588. Doi:10.1093/tas/txy132
- Campos, R. V., Cobuci, J. A., Kern, E. L., Costa, C. N., & McManus, C. M. (2015). Genetic parameters for linear type traits and milk, fat, and protein production in Holstein cows in Brazil. *Asian-Australasian Journal of Animal Sciences*, 28(4), 476–484. Doi: 10.5713/ajas.14.0288
- Chollet, François. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. arXiv. Doi:10.48550/arXiv.1610.02357
- Cominotte, A., Fernandes, A. F. A., Dorea, J. R. R., Rosa, G. J. M., Ladeira, M. M., van Cleef, E. H. C. B., Pereira, G. L., Baldassini, W. A., & Machado Neto, O. R. (2020). Automated computer vision system to predict body weight and average daily gain in beef cattle during growing and finishing phases. *Livestock Science*, 232, 103904. Doi: 10.1016/J.LIVSCI.2019.103904
- Cross, H., Gilliland, D., Durland, P., & Seideman, S. (1983). Beef carcass evaluation by use of a video image analysis system. *Journal of Animal Sci*, 57(4), 908–917. Doi: 10.2527/jas1983.574908x
- Daetwyler, H. D., Villanueva, B., Bijma, P., & Woolliams, J. A. (2008). Inbreeding in genome-wide selection. *Journal of Animal Breeding and Genetics*, 124(6), 369–376. Doi: 10.1111/j.1439-0388.2007.00693.x
- Doeschl, A. B., Green, D. M., Whittemore, C. T., Schofield, C. P., Fisher, A. V., & Knap, P. W. (2004). The relationship between the body shape of living pigs and their carcass morphology and composition. *Animal Science*, 79(1), 73–83. Doi: 10.1017/s1357729800054540
- Fernandes, Arthur F A, Dórea, J. R. R., Fitzgerald, R., Herring, W., & Rosa, G. J. M. (2019). A novel automated system to acquire biometric and morphological measurements and predict body weight of pigs via 3D computer vision 1. *J. Anim. Sci*, 97, 496–508. Doi:10.1093/jas/sky418

- Fernandes, Arthur F A, Dórea, J. R. R., Valente, B. D., Fitzgerald, R., Herring, W., & Rosa, G. J. M. (2020). Comparison of data analytics strategies in computer vision systems to predict pig body composition traits from 3D images. *Journal of Animal Science*, 98(8). Doi: 10.1093/jas/skaa250
- Fernandes, Arthur Francisco Araújo, Dórea, R. J. R., & Rosa, G. J. D. M. (2020). Image Analysis and Computer Vision Applications in Animal Sciences: An Overview. *Frontiers in Veterinary Science*, 7(October), 1–18. Doi: 10.3389/fvets.2020.551269
- Forni, S., Aguilar, I., & Misztal, I. (2011). Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genetics Selection Evolution*, 43(1), 1–7. Doi:10.1186/1297-9686-43-1
- Getu, A., & Misganaw, G. (2015). The Role of Conformational Traits on Dairy Cattle Production and Their Longevities. *OALib*, 02(03), 1–9. Doi:10.4236/oalib.1101342
- Gocheva-ilieva, S., & Yordanova, A. (2022). Predicting the 305-Day Milk Yield of Holstein-Friesian Cows Depending on the Conformation Traits and Farm Using Simplified Selective Ensembles. *Mathematics*. Doi:10.3390/math10081254 Academic
- Gorgulu, O. (2012). Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. *South African Journal of Animal Science*, 42(3). Doi:10.4314/sajas.v42i3.10
- Habier, D., Fernando, R. L., & Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, 177(4), 2389–2397. Doi:10.1534/GENETICS.107.081190
- Halachmi, I., Klopčič, M., Polak, P., Roberts, D. J., & Bewley, J. M. (2013). Automatic assessment of dairy cattle body condition score using thermal imaging. *Computers and Electronics in Agriculture*, 99, 35–40. Doi:10.1016/J.COMPAG.2013.08.012
- Harris, R. (2015). Phenotypic correlations between linear type conformation traits, production and fertility in a once-a-day milked dairy cattle herd. Thesis retrieved 5 May, 2021 from the Massey university website: https://mro.massey.ac.nz/bitstream/handle/10179/7430/02_whole.pdf?sequence=2&isAllowed=y.
- Hassaballah, M., & Awad, A. I. (2020). Deep Learning in Computer Vision. In *Deep Learning in Computer Vision* (Issue March). Doi:10.1201/9781351003827
- Hovinen, M., Siivonen, J., Taponen, ‡ S, Hänninen, L., Pastell, † M, Aisla, A.-M., & Pyörälä,

- S. (2008). Detection of Clinical Mastitis with the Help of a Thermal Camera. *Journal of Dairy Science*, 91, 4592–4598. Doi:10.3168/jds.2008-1218
- International Committee of Animal Recording. (2015). 1 . Conformation Recording of Dairy Cattle. June, 1–42. retrieved 1 August, 2021 from the ICAR website: <https://www.icar.org/wp-content/uploads/2015/08/Conformation-Recording-CR-WG.pdf>
- Janković, D., Trivunović, S., & Samolovac, L. (2020). Characteristics of linear traits of udder and angularity in Holstein- friesian cows and their correlation with milk yield traits. *36(4)*, 407–416. Doi:10.2298/BAH2004407D
- Jensen, D. B., van der Voort, M., & Hogeveen, H. (2018). Dynamic forecasting of individual cow milk yield in automatic milking systems. *Journal of Dairy Science*, 101(11), 10428–10439. Doi:10.3168/JDS.2017-14134
- Jeretina, J., Babnik, D., & Škorjanc, D. (2016). Prediction of Standard Lactation Curves for Primiparous Holstein Cows by Using Corrected Regression Models Prediction of standard lactation curves for primiparous Holstein cows by using corrected regression models. *Italian Journal of Animal Science*. Doi: 10.4081/ijas.2015.3776
- Jingar, S., Mehla, R. K., Singh, M., & Roy, A. K. (2014). Lactation Curve Pattern and Prediction of Milk Production Performance in Crossbred Cows. *Journal of Veterinary Medicine*, 2014, 1–6. [Doi:10.1155/2014/814768](https://doi.org/10.1155/2014/814768)
- Justin Johnson (2019). Introduction to deep learning for computer vision. Lecture retrieved 5 January 2022 from youtube: <https://www.youtube.com/watch?v=dJYGatp4SvA>
- Kashiha, M., Bahr, C., Haredasht, S. A., Ott, S., Moons, C. P. H., Niewold, T. A., Ödberg, F. O., & Berckmans, D. (2013). The automatic monitoring of pigs water use by cameras. *Computers and Electronics in Agriculture*, 90, 164–169. Doi:10.1016/J.COMPAG.2012.09.015
- Kern, E. L., Cobuci, J. A., Costa, C. N., McManus, C. M., Campos, G. S., & Almeida, T. P. (2015). Genetic association between longevity and linear type traits of holstein cows. *Scientia Agricola*, 72(3), 203–209. Doi: 10.1590/0103-9016-2014-0007
- Khan, M. S. M. A., & Khan, M. S. M. A. (2016). Genetic and phenotypic correlations between linear type traits and milk yield in Sahiwal cows. *Pakistan Journal of Agricultural Sciences*, 53(2), 483–489. Doi:10.21162/PAKJAS/16.3369
- Khoshnoudi-Nia, S., & Moosavi-Nasab, M. (2019). Prediction of various freshness indicators

- in fish fillets by one multispectral imaging system. *Scientific Reports*, 9(1).
Doi:10.1038/S41598-019-51264-Z
- Kliś, P., Piwczyński, D., Sawa, A., & Sitkowska, B. (2021). Prediction of lactational milk yield of cows based on data recorded by AMS during the periparturient period. *Animals*, 11(2), 1–11. Doi:10.3390/ani11020383
- Kongsro, J. (2014). Estimation of pig weight using a Microsoft Kinect prototype imaging system. *Computers and Electronics in Agriculture*, 109, 32–35. Doi:10.1016/J.COMPAG.2014.08.008
- Li, N., Ren, Z., Li, D., & Zeng, L. (2020). Review: Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming. *Animal*, 14(3), 617–625. Doi:10.1017/S1751731119002155
- Liseune, A., Salamone, M., Van den Poel, D., Van Ranst, B., & Hostens, M. (2020). Leveraging latent representations for milk yield prediction and interpolation using deep learning. *Computers and Electronics in Agriculture*, 175(July), 105600. doi:10.1016/j.compag.2020.105600
- Lucas, D., Brun, A., Gispert, M., Carabús, A., Soler, J., Tibau, J., & Font-i-Furnols, M. (2017). Relationship between pig carcass characteristics measured in live pigs or carcasses with Piglog, Fat-o-Meat'er and computed tomography. *Livestock Science*, 197, 88–95. Doi:10.1016/J.LIVSCI.2017.01.010
- Luis Nunes, J., Piquerez, M., Pujadas, L., Armstrong, E., Fernández, A., & Lecumberry, F. (2015). Beef quality parameters estimation using ultrasound and color images. doi:10.1186/1471-2105-16-S4-S6
- Martins, R. F. S., do Prado Paim, T., de Abreu Cardoso, C., Stéfano Lima Dallago, B., de Melo, C. B., Louvandini, H., & McManus, C. (2013). Mastitis detection in sheep by infrared thermography. *Research in Veterinary Science*, 94(3), 722–724. Doi:10.1016/J.RVSC.2012.10.021
- Matthews, S. G., Miller, A. L., Plötz, T., & Kyriazakis, I. (2017). Automated tracking to measure behavioural changes in pigs for health and welfare monitoring. *Scientific Reports*, 7(1), 1–12. Doi:10.1038/s41598-017-17451-6
- Metzner, M., Sauter-Louis, C., Seemueller, A., Petzl, W., & Zerbe, H. (2015). Infrared thermography of the udder after experimentally induced *Escherichia coli* mastitis in cows. *The Veterinary Journal*, 204(3), 360–362. Doi:10.1016/J.TVJL.2015.04.013

- Nasirahmadi, A., Edwards, S. A., & Sturm, B. (2017). Implementation of machine vision for detecting behaviour of cattle and pigs. *Livestock Science*, 202, 25–38. Doi:10.1016/J.LIVSCI.2017.05.014
- Nguyen, Q. T., Fouchereau, R., Frénod, E., Gerard, C., & Sincholle, V. (2020). Comparison of forecast models of production of dairy cows combining animal and diet parameters. *Computers and Electronics in Agriculture*, 170. Doi:10.1016/j.compag.2020.105258
- O’Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G. V., Krpalkova, L., Riordan, D., & Walsh, J. (2020). Deep Learning vs. Traditional Computer Vision. *Advances in Intelligent Systems and Computing*, 943(Cv), 128–144. Doi:10.1007/978-3-030-17795-9_10
- Ogundeji, A. A., Lakew, H., Tesfahuney, W., & Lombard, W. (2021). Influence of heat stress on milk production and its financial implications in semi-arid areas of South Africa. *Heliyon*, 7(5), e06202. Doi:10.32964/tj19.5
- Ozkaya, S. (2015). Prediction possibility of milk yield from udder measurements using digital image analysis on Holstein cows. *Indian Journal of Animal Research*, 49(3), 388–391. Doi:10.5958/0976-0555.2015.00050.3
- Paluchowski, L. A., Misimi, E., Grimsmo, L., & Randeberg, L. L. (2016). Towards automated sorting of Atlantic cod (*Gadus morhua*) roe, milt, and liver – Spectral characterization and classification using visible and near-infrared hyperspectral imaging. *Food Control*, 62, 337–345. doi:10.1016/J.FOODCONT.2015.11.004
- Parke, R. et al. (1999) ‘Genetic and phenotypic parameter estimates between production, feed intake, feed efficiency, body weight and linear type traits in first lactation Holsteins’, *Canadian Journal of Animal Science*, 79(4), pp. 425–431. doi: 10.4141/a99-008.
- Pezzuolo, A., Guarino, M., Sartori, L., González, L. A., & Marinello, F. (2018). On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera. *Computers and Electronics in Agriculture*, 148, 29–36. Doi:10.1016/J.COMPAG.2018.03.003
- Phahlane, H., Mthembeka, Z. A., Lekganyane, M. S., Mofolo, K. M., Gininda, P., Ramonyai, D., & Lubbe, P. (2021). Quarterly economic overview. The Department of Agriculture, Forestry and Fisheries, Volume 19, 5–24. Retrieved 2 February 2022 from the Department of Agriculture, Land reform and Rural Development website: <https://www.dalrrd.gov.za/Portals/0/Statistics%20and%20Economic%20Analysis/Economic%20Analysis/Quarterly%20Economic%20Overview%20of%20the%20%20Agriculture%20Forestry%20%20Fisheries%20%20Volume%2019%20Number%201%20First>

[%20Quater%202021.doc.pdf](#).

- Psota, E. T., Mittek, M., Pérez, L. C., Schmidt, T., & Mote, B. (2019). Multi-pig part detection and association with a fully-convolutional network. *Sensors (Switzerland)*, 19(4), 1–24. Doi:10.3390/s19040852
- Saberioon, M., Gholizadeh, A., Cisar, P., Pautsina, A., & Urban, J. (2016). Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues. Doi:10.1111/raq.12143
- Sassi, N. Ben, Averós, X., Estevez, I., Nicol, C., & Rodenburg, T. B. (2016). *animals Technology and Poultry Welfare*. Doi:10.3390/ani6100062
- Smith, G. (2021). South Africa’s dairy sector strives to shake off its past. In *The Daily Churn*. Darigold. Retrieved 3 August 2021: <https://www.darigold.com/south-africas-dairy-sector-strives-to-shake-off-its-past/>
- South African Milk Processors’ Organization. (2021). Summary of Key Market Signals for the Dairy Industry , August 2021 Edition. Annual Report, August 2021. Retrieved 5 March 2022 from the South African Milk Processors’ Organization website: <https://sampro.co.za/wp-content/uploads/2021/09/August-2021-Edition-of-Summary-of-Key-Market-Signals-for-the-Dairy-Industry.pdf>
- 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. Doi:10.3168/JDS.2015-10607
- Stamschror, J., & Card, S. (2000). *Judging Dairy Cattle*. System, July, 1–8. Retrieved 12 May 2021: <https://www.yumpu.com/en/document/read/11309884/judging-dairy-cattle-usda-aphis>
- Tapki, I. (2013). Genetic and Phenotypic Correlations between Linear Type Traits and Milk Production Yields of Turkish Holstein Dairy Cows By. *Green Journal of Agricultural Sciences*, 3(11), 755–761. Doi:10.15580/GJAS.2013.11.072913763
- Vallejo, R. L., Leeds, T. D., Gao, G., Parsons, J. E., Martin, K. E., Evenhuis, J. P., Fragomeni, B. O., Wiens, G. D., & Palti, Y. (2017). Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genetics Selection Evolution*, 49(1), 1–13. [doi:10.1186/s12711-017-0293-6](https://doi.org/10.1186/s12711-017-0293-6)
- Visser, C. et al. (2020) ‘Phenomics for sustainable production in the South African dairy and beef cattle industry’, 10(2). doi:10.1093/af/vfaa003.

- Waal, H. De, & Blom, C. (2020). Annual Milk Cattle Bulletin. National Milk Recording and Improvement Scheme. Retrieved 14 July 2021 from the Agricultural Research Council website:
<https://www.arc.agric.za/ARC%20Newsletters/Annual%20Milk%20Cattle%20Bulletin.%20Issue%2023,%202020.pdf>.
- Wall, E., White, I. M. S., Coffey, M. P., & Brotherstone, S. (2005). The relationship between fertility, rump angle, and selected type information in Holstein-Friesian cows. *Journal of Dairy Science*, 88(4), 1521–1528. doi:10.3168/jds.S0022-0302(05)72821-6
- Wang, X., Zhao, M., Ju, R., Song, Q., Hua, D., Wang, C., & Chen, T. (2013). Visualizing quantitatively the freshness of intact fresh pork using acousto-optical tunable filter-based visible/near-infrared spectral imagery. *Computers and Electronics in Agriculture*, 99, 41–53. Doi:10.1016/j.compag.2013.08.025
- Wassenberg, R., Allen, D., & Kemp, K. (1986). Video image analysis prediction of total kilograms and percent primal lean and fat yield of beef carcasses. *Journal of Animal Sci*, 62(6), 1609–1616. Doi:10.2527/jas1986.6261609x
- Yakubu, A. (2011). Path analysis of conformation traits and milk yield of Bunaji cows in smallholder ' s herds in Nigeria. *Agricultura Tropica et Subtropica*, 44(3), 152–157.
- Zapotoczny, P., Szczypiński, P. M., & Daszkiewicz, T. (2016). Evaluation of the quality of cold meats by computer-assisted image analysis. *LWT - Food Science and Technology*, 67, 37–49. Doi:10.1016/J.LWT.2015.11.042
- Zhang, M., Luo, H., Xu, L., Shi, Y., Zhou, J., Wang, D., Zhang, X., Huang, X., & Wang, Y. (2022). Genomic Selection for Milk Production Traits in Xinjiang Brown Cattle. *Animals*, 12(2), 1–13. Doi:10.3390/ani12020136

CHAPTER THREE: Predicting milk yield using a cow's rear-view image through computer vision deep learning

Abstract

The current study was conducted to establish the predictability of cows' 305-day milk yield using cow's rear-view images through deep learning. A total of 1238 rear-view images from 743 Holstein cows within their first or second parity and their corresponding first lactation milk yield values were split into the training and testing data at the ratio of 80:20, respectively. The training data was augmented four times more, and the test data was left unaugmented. Augmentation increased the training data to 5005 images. The training images, both augmented and unaugmented images, were again split into training and validation data at the ratio of 80:20, respectively. Based on these rear-view images and their corresponding milk yield values, a deep learning model was trained and validated for milk yield prediction. Specifically, an Xception convolutional Neural Network architecture was the model used. The established model was then tested using the test data. There was a drop in performance from the model validation results and the model test results: i.e. from an MAE value of 373.6 kg, an R^2 value of 0.90 and an RMSE of 615.9 kg on the validation data to an MAE value of 1148.3 kg, the R^2 value of 0.30 and RMSE of 1480 kg on the testing data. This was because some validation images were augmented; hence they were not entirely new cases for the model. The poor predictions observed based on the test results were attributed to variations in the fullness of the udder with milk in the images, full automation of the prediction process, differences in the camera's distance from the cow and posture differences, especially regarding the rear leg position. It was concluded that milk yield can be predicted from rear-view images when using computer vision deep learning.

Keywords: Milk Yield, deep learning, rear-view images, linear-type traits, Xception network

3.1. Introduction

Linear-type traits explain various performance characteristics in cattle, such as fertility (Zindove et al., 2014; Harris, 2015), longevity (Kern et al., 2015) and milk production (Khan & Khan, 2016; Gocheva-ilieva & Yordanova, 2022). Traits including angularity, stature, body depth, rump width, chest width, udder depth, rear udder width, fore udder attachment and rear udder height are used to explain milk yield variability. More image views would be required when using image data to make predictions from these traits, as all these traits cannot be accommodated in one picture. For example, it is nearly impossible to have a perfect view of

the fore udder attachment and rear udder width in the same picture. Some traits can only be viewed from the back of the cow, and some from the side.

There is a myriad of research on the correlation between linear-type traits and milk yield (Janković et al., 2020; Getu & Misganaw, 2015; Tapki, 2013). However, only a handful address the actual milk yield prediction based on these correlations. The few pieces of research on milk yield prediction from conformation traits often concentrate on only rear-view traits. This is possibly because of a more accessible view for the observers when the cows are in restrainers or milking parlour. Yakubu (2011) examined the predictability of milk yield from nine linear-type traits, and only heart girth was not a rear-view trait. Ozkaya (2015) examined the predictability of milk yield based on udder measurement, all seen from a cow's rear-view. This study, therefore, seeks to evaluate whether deep learning can make milk yield prediction better when using rear-view images

Gocheva-ilieva & Yordanova (2022) examined the predictability of milk yield from stature, chest width, rump width, rear legs rear-view, rear leg set side-view, hock development, bone structure, foot angle, foot depth and udder width. Among the linear-type traits examined, Gocheva-ilieva and Yordanova (2022) found udder width, chest width, and stature to be the three essential conformation traits in predicting milk yield. These are traits one will never find on the same image view and are not all udder traits. The list of evaluated traits by Gocheva-ilieva and Yordanova (2022) may not be exhaustive, and more traits correlated to milk yields were not evaluated, such as angularity and rear udder height (Campos et al., 2015). Including more linear-type traits in the milk yield prediction could improve the accuracy of prediction.

Ozkaya (2015) established that MY predictions were better when all the rear udder traits were considered as gestalt than when they were used individually. The current study goes a step further to evaluate the predictability of milk yield when all the possible traits seen from the posterior of a cow are considered.

3.2. Materials and Methods

3.2.1. Description of the study sites

Data were collected from four dairy farms in and around Humansdorp in the Eastern Cape, South Africa. Humansdorp has a prevalent semi-arid climate. It is characterised by long, cool, windy winters; and short and warm summers. It is dry and mostly clear all year round. Over the year, the temperature typically varies from 9.4°C to 26.1°C and is rarely below 6.7°C or above 29.4°C. For this report, the geographical coordinates of Humansdorp are 34.0027° S, 24.7440° E, and 36,58 metres elevation (World weather online. 2022).

The warm season persists for approximately three months, from mid-December to mid-March, with an average daily high temperature above 24.4°C. The hottest month of the year is February, with an average high of 25.6°C and a low of 18.3°C. On the other hand, the cool season lasts for approximately 4.1 months, from the end of May to the beginning of October, with an average daily high temperature below 20.6°C. The coldest month of the year is July, with an average low of 10°C and a high of 19.4°C (World weather online. 2022).

There is no marked seasonal variation in the frequency of wet days (i.e., those with greater than 1.02 mm of liquid precipitation) in Humansdorp. The average frequency of wet days is 14 % but ranges from 10 to 18 %. Humansdorp experiences some seasonal variation in monthly rainfall. Rain falls throughout the year, and November receives the most rain, with an average rainfall of 38.1. The month with the least precipitation is May, with an average of 20.3 mm (World weather online. 2022).

The wind experienced at a particular location is highly reliant on local topography. The windier part of the year lasts for approximately seven and half months, from early June to mid-January. During this time, the average wind speed is more than 15.9 km per hour. The windiest month of the year is October, with an average hourly wind speed of 17.4 km per hour. The calmer time of the year lasts for roughly 4.6 months, from mid-January to early June. March is the calmest month of the year, with an average hourly wind speed of 14.7 km per hour. The wind is often from the west from the beginning of April to mid-October, with a peak percentage of 51 % around June 4. The wind is most often from the east for 5.8 months, from mid-October to early April, with a peak percentage of 38 % in January (World weather online. 2022).

3.2.2. Study variables

The dependent variable in this study was milk yield. A cow's rear-view image was the independent variable. The number of milkings, feeding system and image dimensions were controlled in this study. Variations in the number of milkings and the feeding system have a bearing on the amount of milk a cow can produce. Changes in image dimensions could result in distortion of the image proportions, hence interfering with the learning process.

3.2.3. The data

All the cows used for this research were reared under the pasture-based system. They were all milked twice a day. The first milking commenced either at 0300h or 0400h and the last milking either at 1400h or 1500h, depending on the farm.

There is no rule of thumb on the minimum required sample size, especially when dealing with continuous variables for deep learning. For a confidence interval greater than 0.1, a minimum

dataset size of 1000 was required (Blatchford et al., 2021). For the current study, a data size of 1238 rear-view images was used. Three hundred sixty rear-view images were from farm A, 260 were from farm B, 394 were from farm G, and 224 rear-view images were from farm M. Some of the photos were from the same cow but viewed differently; therefore, all the pictures were different. The actual number of cows used in the analysis was 743. Two hundred and twenty eight cows were from farm A, 151 from farm B, 241 from farm G and 123 from farm M. Table 3.1 below is a summary of the cows and the images used in this study.

Table 3.1: Description of the data

Farm	number of cows	number of images
A	228	360
B	151	260
G	241	394
M	123	224
Total	743	1238

The relative representation of each of the farms i.e. the percentage of images relative to number of cows in the dataset ranged from 63% (228/360) to a low of 55% (123/224). The effect of the farm variations in image ratios to the number of cows on the prediction process are expected to be insignificant. This is because the training data were taken from farms with the same feeding and milking systems and within the same climatic region. Variations in the average milk yield from one farm to the other can be attributed to different breeding programs, which also manifests in the cow's physical features. Also, the over-representation of data can only improve the model but does not compromise on the less represented data.

The data consisted of cows' rear-view images and the cows' corresponding 305-day milk yield values from the first lactation. The pictures taken were of Holstein cows within their first or second lactation. These were the cows born between 2017, 2018 and early 2019. Milk yield data was prospective for cows within their first lactation, yet it was retrospective for cows within their second lactation. Milk yield prediction is pertinent earlier in a cow's life, hence the use of first lactation milk yield. Photos were from cows within the body condition score (BCS) range of 3.25 to 3.75, which is the farmers' desired BCS range. Information on the body

condition scores was available in farm operations software. Selection of a BCS range was done so as to filter out most malnourished and diseased cows.

The desired rear-view features were rump width, udder depth, udder height and udder width. It is salient to note that since we were using cow visual images, some features not named above may have impacted the overall milk yield prediction.

All the images used were taken from the end of September to the beginning of November. The cow images were taken as the cows were in the pastures. This was because there is minimal interference in the farm operations when the cows are in the fields, and the cows are more relaxed, making it easier to take pictures. Since the rear-view images were not captured in the chronology of the cow identities, cow identities were written down in the order of photo capturing. The farm name was written at the beginning of each page to differentiate data from different farms. This also made reference easier as some of the cows from different farms had the same identity.

All the images were taken during the day, between 0700h and 1730h. Time, regarding milking, was not considered in the image-capturing process. There is no variation, within a day, on the standard linear-type traits. Needless to say that the udder is visibly swollen just before milking and deflated soon after milking. However, this is not expected to result in temporary changes in the rear udder width since it is evaluated at the point of udder attachment (ICAR, 2015).

The milk yield records for each cow and parity, BCS information and lactation number information were then extracted from the farm operations recording software.

3.2.4. Data pre-processing

Data pre-processing involved image renaming, image editing, removal of images of cows that are out of range and, finally, preparation of CSV files with image names and their corresponding milk yield values. The first step was to rename the images from the autogenerated image names to the cow identities to correspond to the cow in the picture.

3.2.4.1. The image naming system

Each image name contained an R at the end of the identity to indicate that the image is a rear-view image. Also, some cows had more than one image for the same view. The permutations would be denoted by “(X)” after the image name: where X is the permutation number. For example, “18123S (2)” would be the second side-view image of the cow 18123.

Some cows from different farms had the same image name; hence, there was a need to differentiate between these cows. A letter was put at the beginning of the image ID to separate images from different farms. For example, image “G18123R (3)” would be the third, signified by “(3)”, rear-view image, indicated by R, of the cow 18123 from Glen-herd farm, denoted by G. However, farm A had no symbol at the beginning of its image names. Farm B images had B at the beginning of their image names; the letter G was used to identify farm G, and an M at the beginning of the image names identified farm M images.

The images that were created after image augmentation have unique identifiers. Firstly, they all have an underscore at the beginning of the image name, so they are sequentially separated from the unaugmented files. Each image has four augmentations; therefore, each has aug_1 to aug_4 written at the end of the actual image name. For example, augmentation two of the image G18123R has the following image name: G18123R_aug_2.

3.2.4.2. Image editing

Some images contained more than one cow; this had the possibility of interfering with the prediction analysis. Other cows appearing in the picture were shaded out to preclude interference, leaving just one cow for each image. The dimensions of the images, i.e. 3872 x 2592, were not altered in the editing process. Figure 3.1 below is an example to show how the shedding out of other cows was done:



Figure 3.1: Image before shedding out undesired cows(left); image after removing undesired cows(right)

Later, when all the photos from different farms were combined, and the two CSV files were created, the training and validation (TV) samples images were augmented to generate more samples for analysis. It is salient to note that the testing data, different from the validation data, was not augmented. This was so that no single permuted image from the TV sample reappears in the test data, resulting in the model falsely performing well. However, an analysis was done where all the data was augmented, and the test data with 20% of all the data created was randomly drawn.

Four random selections from three augmentation methods were used for each picture. This means that one of the three augmentation methods was reused; however, it did not produce the same image as one out of the four augmented images would be a combination of two augmentations. The augmentation methods included were random noise, 25-degree random rotation either to the left or the right and a horizontal flip. This means that for each image used, there were five permutations: the original image plus four augmented images.

As much as image proportions were uniformly distorted when the photos were loaded into the model, these augmentation types were selected because they do not distort image proportions. The research, hypothetically, deals with measurement traits: widths and heights. Therefore, distortion of image proportions at the augmentation stage could reduce accuracy.

3.2.4.3. Selection of pictures and CSV files creation

The edited images from all four farms were put into one folder. The names for each image were then extracted and put on a spreadsheet as they appeared on the image. Annexed to the rear-view images column was the 305-day milk yield column.

Not all images captured were used for the analysis; some of the cows had no 305-day milk yield records, some were much older than 56 months, even though they were within the desired parity or lactation number, and some were out of parity or lactation number range. For the aforementioned reason, 524 rear-view images were filtered out, leaving out 1238 usable images. The CSV file was then updated only to contain the names of the pictures used for the analysis and their corresponding 305-day milk yield values.

Twenty percent of the data were randomly extracted and pasted onto a separate CSV file from the initial CSV file to make the testing data. The remaining 80 % was then used to make the CSV file for the TV data. The TV images were augmented using the training and validation

CSV information. The augmented file names and their corresponding milk yields were then annexed to the rows on the TV CSV file. The TV CSV now contained both the TV augmented data and TV unaugmented data. It is important to note that the validation data was a randomly selected 20 % of the TV sample after augmentation. Random splitting of the validation data from the training data was put on the model command; hence the CSV file did not require any splitting to separate the training data and the validation data. All the split ratios used in this research were based on a recommendation by (Majurski, 2019).

3.2.5.Data analysis

Data analysis was split into hyperparameter optimisation and the actual tests using the optimal parameters. All the analyses were carried out using the Xception deep learning network/model (Chollet, 2017). The network was implemented using the TensorFlow framework. Due to inadequate graphics processing unit (GPU) space, the models were trained on a Central Processing Unit (CPU) allocator. The two columns in the CSV files were used to call the images and their corresponding milk yield values.

The Mean Square Error (MSE), mean absolute error (MAE), and the coefficient of determination (R^2) values were used to evaluate and update trait weights. The root mean square error (RMSE), MAE, Mean Absolute Percentage Error (MAPE) and R^2 values were used to report the research findings on the test data. The following are the equations for each of these evaluation techniques:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Where, for the i th record:

y = observed value

\hat{y} = predicted value

n =total number of data points.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Where the sum of squared residual has the following formula:

$$SS_{res} = \sum_i (y_i - \hat{y}_i)^2 = \sum_i e_i^2$$

And the sum of squared total has the following formula:

$$SS_{tot} = \sum_i (y_i - \bar{y})^2$$

For the i th value of the SS_{res} and SS_{tot} :

y = observed values

\hat{y} = predicted value

$e = residuals = y_i - \hat{y}_i$

\bar{y} = mean of the observed data:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

3.2.5.1. Hyperparameter optimisation and the optimum parameters used

Hyperparameter optimisation or parameter tuning is the process of establishing a set of optimal hyperparameters for a learning algorithm in machine learning. A hyperparameter is a parameter whose value regulates the learning process. This process was run after image augmentation. Basically, parameter tuning is a process of trial and error, running the same experiment but on different parameters until one finds the parameters that give better results.

Parameter optimization was done after data augmentation to minimize parameter changes that may be brought about by augmentation. Mainly the validation MAE results were used to determine the best parameters. The following parameters were examined, the best learning rate, whether or not to use a learning rate decay, whether or not to use dropouts, and the optimum number of epochs. A learning rate decay/schedule is a predefined structure that regulates the learning rate between epochs or iterations as the learning progresses. The number of epochs is the number of complete passes on all the training data. Dropout is a case where some features on the image are randomly left out during model training.

Due to lower CPU memory, the batch size could not be raised above 32. Therefore, the batch size for all the trials was kept at 32. Batch size is the number of needed samples before the model is updated. According to Kandel and Castelli (2020), a batch size of 32 is a good default value. In the same paper, Kandel and Castelli (2020) also mentioned that batch size's significant impact is on training time and not on performance. The larger the batch size, the faster the computation; however, more memory is also required.

One extremely pertinent hyperparameter when it comes to model performance is the learning rate. It significantly affects the learning process. There is a significant correlation between learning rate and batch size. Lower learning rates perform better with smaller batch sizes and vice versa (Kandel & Castelli, 2020). However, Kandel and Castelli (2020) recommended using a smaller batch size, e.g., 32 or 64, with a lower learning rate, to using a higher learning rate with a bigger batch size.

Learning rates between 0.0004 and 0.001 were tried for the rear-view analysis model. Based on the MAE, the optimum learning rate was 0.0006. The LR comparison results were done at 38 epochs with an LR scheduler and no dropout. All the evaluation metrics were in harmony that a learning rate of 0.0006 gave the best results. Table 3.2 shows a comprehensive report of the comparison between various learning rates. Below table 3.2 is figure 3.2, plotted based on the validation MAE from table 3.2, clearly showing the LR with the lowest MAE.

Table 3.2: Rear-view learning rate comparison results at 38 epochs, with LR scheduler and no dropout.

Training					Validation			
LR	MSE	RMSE	MAE	R ²	MSE	RMSE	MAE	R ²
0,0004	390266,558	624,713	598,451	0,897	655091,806	809,377	680,055	0,822
0,0005	39733,562	199,333	152,703	0,910	405158,008	636,520	403,965	0,810
0,0006	31381,393	177,148	134,593	0,992	379376,848	615,936	373,632	0,897
0,0007	60643,801	246,210	187,205	0,984	458337,402	677,006	427,388	0,875
0,0008	108727,229	329,738	258,941	0,971	606733,920	778,931	525,694	0,835
0,0009	55489,657	235,562	177,742	0,985	662253,007	813,790	490,997	0,820
0,001	158276,694	397,840	325,803	0,958	709818,231	842,507	602,033	0,807

LR= Learning rate, MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error, R²= coefficient of determination.

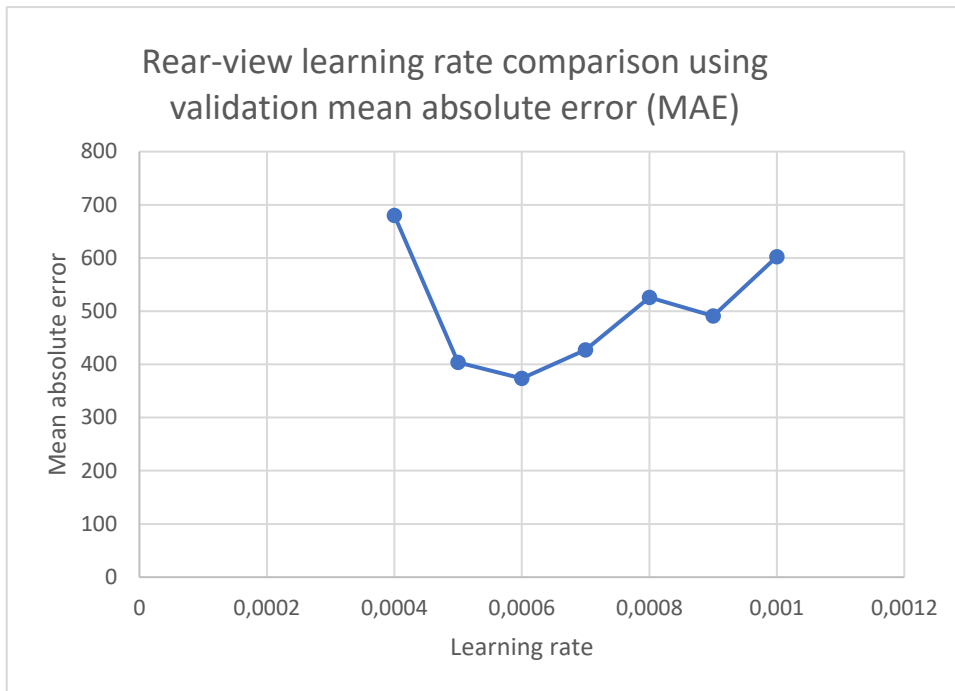


Figure 3.2: Rear-view learning rate comparison results at 38 epochs, with LR scheduler and no dropout.

The learning rate was used together with a learning rate schedule. The learning rate schedule used was a time-based decay rate. A learning rate decay is used because choosing a large initial learning rate and then decreasing it over time often results in better converged and more well-performing models. Early in model training, the model makes giant steps towards the gradient space, and a large learning rate helps to quickly find the coarse values. In the later stages of model training, the opposite is true. The model already has approximately the correct gradients, and it just requires a slight extra push to find the last few percentage points. At this point, a large gradient is now inappropriate because it will overshoot the point of optimality. Instead of converging on the global minima, the model will bounce around it; hence a learning rate decay schedule is required (Andrew et al., 2013). The following formula was used to calculate the learning rate decay rate:

$$\text{Learning Rate decay} = \frac{LR * 1}{(1 + 10 * decay * epoch)}$$

Where:

LR= previous learning rate

Epoch= the current epoch number

$$\text{Decay} = \frac{\text{initial learning rate}}{\text{Epochs}}$$

The optimiser used in this research is adaptive, i.e. the Adam optimiser, which has a learning rate schedule embedded in it. For this reason, literature conflicts on whether a learning rate scheduler should be used together with the Adam optimiser. We then compared the results from one with a learning rate schedule and the other without the scheduler.

Tests to examine whether or not to use an LR scheduler were done using the optimum Learning rate of 0.0006 and 38 epochs. Based on the validation for all evaluation metrics, better results were found when the LR scheduler was present. Table 3.3 below shows the results for comparing a model with an LR scheduler to one without a scheduler.

Table 3.3: Comparison between a model with a learning rate scheduler and one with no LR scheduler.

Evaluation matrix	LR decay present	LR decay absent
MSE	379376,847	452528,749
RMSE	615,936	672,703
MAE	373,632	413,193
R ²	0,897	0,877

LR= Learning rate, MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error, R²= coefficient of determination.

The Adam optimiser was used because it is currently the best performing optimiser in image and language processing. Its adaptiveness makes it perform better. It assigns different learning rates to various parameters, quickly learning which parameters are more important in making predictions and which ones are less or not meaningful (Kingma & Ba, 2015).

The number of epochs is the number of complete passes on all the training data; this is different from the batch size, the number of needed samples before the model is updated. Too many epochs beyond the point of convergence may result in data memorisation and model overfitting. Too few epochs before the point of convergence may result in incomplete learning and model underfitting. Therefore, there is a need to select the correct number of epochs, which is just when the validation curve stabilises.

The model was run at fifty epochs. The purpose was to observe the epoch range at which the validation curve stabilises. Having more epochs after the validation error has stabilised is redundant as no more learning occurs beyond the point of stabilisation. About thirty-eight epochs were found to be optimum for the rear-view. Figure 3.3 shows the learning curves for training and validation based on the MAEs. The test was done with the LR of 0.0006 and with a learning rate schedule present.

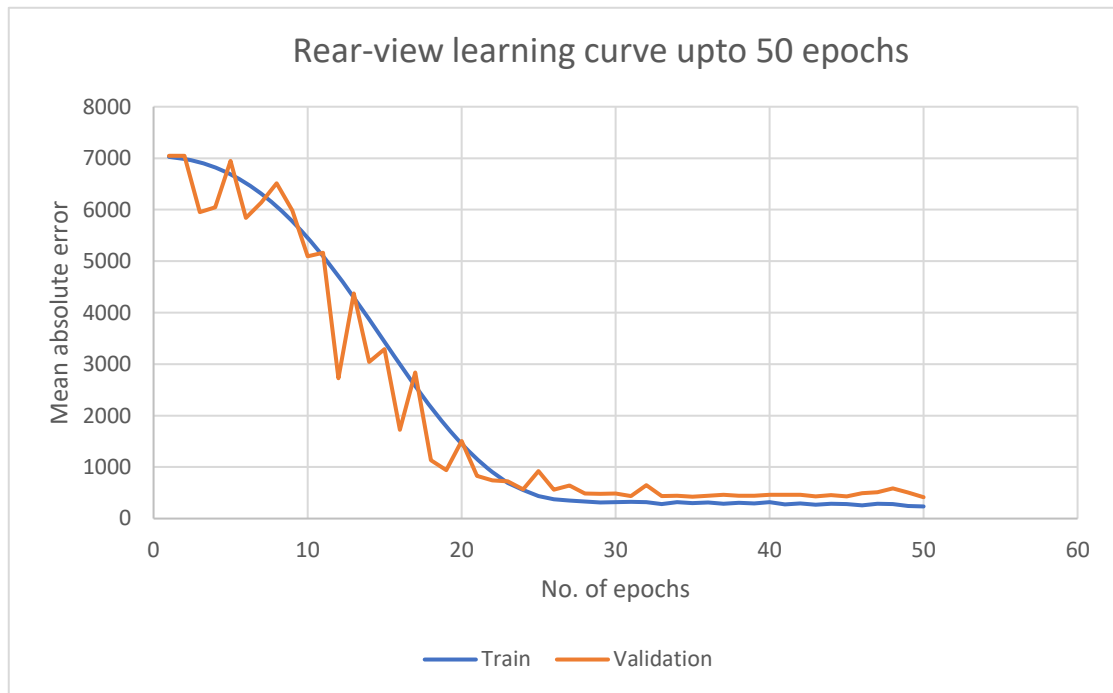


Figure 3.3: Rear-view learning curves for the training and validation.

Dropouts are used to reduce model overfitting or memorisation of data..A model with a dropout rate of 0.4 was compared to those with no dropout. Based on the comparison results, better results were seen in models with no dropouts. The tests were done with an LR of 0.0006 with an LR scheduler. Table 3.4 compares models with dropouts and the ones without dropouts.

Table 3.4: Effects of dropouts (dropout rate=0.4) on model learning.

Evaluation matrix	Dropout present	Dropout absent
MSE	435629,222	379376,847
RMSE	660,022	615,936
MAE	432,319	373,632

R ²	0,881	0,897
----------------	-------	-------

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R²= coefficient of determination.

In summary, the optimiser and batch size used was determined based on a priori knowledge, and the Adam optimizer and a batch size of 32 were used for the final model. Based on empirical evidence, for the rear-view model, dropouts were not used, an initial LR of 0.0006 was used, the model had a time-based scheduler, and 38 epochs were used. Table 3.5 is a summary of the hyperparameters used for the final rear-view model.

Table 3.5: Optimum parameter used for the analyses.

Hyperparameter	Rear-view analysis
Optimiser	Adam
Initial learning rate	0.0006
Learning rate scheduler	Yes
Batch size	32
The final epochs number	38
Drop out	No

3.2.6. Description of the data

Table 3.6 is a description of the data that was used for this research. For training and validation, the sample size was 4987, including augmented images. The average, maximum, minimum and standard deviation in milk yield were 7048.6 kg, 13843 kg, 1644 kg and 1941,9 kg, respectively. The training and validation data were randomly split during the model training after augmentation at the ratio of 80:20, respectively. The testing data constituted a random sample of 20 % of all the data before augmentation. The testing sample size was 237. The

average, maximum, minimum and standard deviation of milk yield were 7111.3 kg, 12789 kg, 2666 kg and 1772,0 kg, respectively.

Table 3.6: Description of the testing and the training and validation data in terms of sample size, average, maximum, minimum and standard deviation of the 305-day milk yield.

	Training and validation	Testing
Number of samples	4987	237
Average MY (kg)	7048,623	7111,338
Maximum MY (kg)	13843	12789
Minimum MY (kg)	1644	2666
Standard Deviation (kg)	1941,929	1772,015

MY= milk yield

3.2.6.1. Frequency distribution for the training and the test datasets.

The testing data had approximately 30 % of its milk yield values between 7000 and 8000kg, whereas the training and validation data had only about 21 % within this range. Figure 3.4a and b below show a full frequency distribution report.

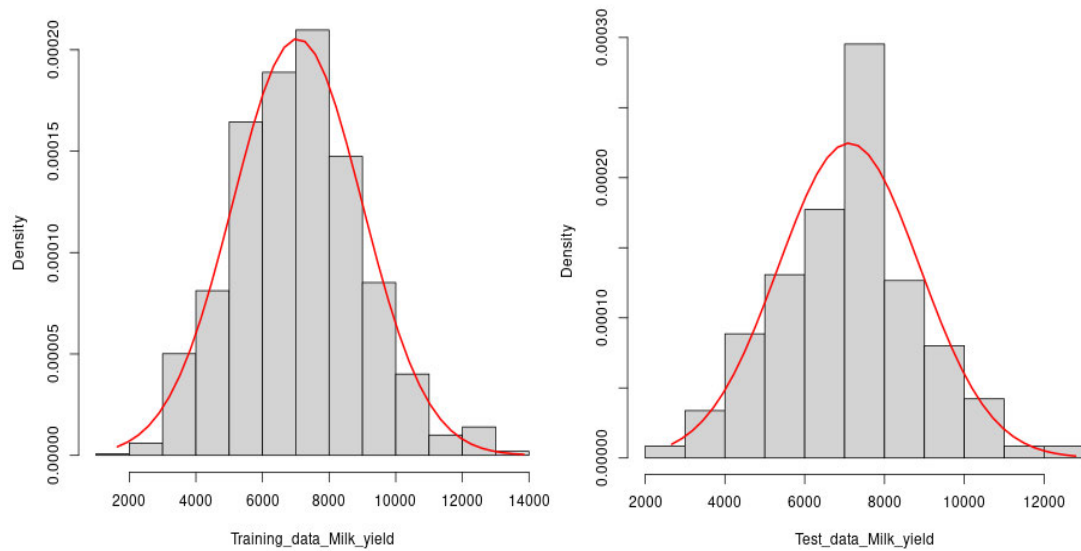


Figure 3.4: 305-day milk yield frequency distribution for the training and validation data: 3.4a.(left) and the testing data: 3.4b. (right).

3.3. Results and Discussion

The established model for the prediction of milk yield from the cows' rear-view images was then run on testing data, which was not used for training or validation and did not contain augmented images. The 305-day milk yield could be predicted using rear-view images. Based on the results in Table 3.7, however, the model's performance drastically dropped from an MAE value of 373.6 kg, an R^2 value of 0.90 and an RMSE of 615.9 kg on the validation data to an MAE value of 1148.3 kg, the R^2 value of 0.30 and RMSE of 1480 kg on the testing data. For the model training, the MSE was 31381.4, the RMSE was 177.1 kg, the MAE was 134.6 kg, and the R^2 was 0.99. Table 3.7 shows the rear-view model's training, validation, and testing results at the established optimum parameters.

Table 3.7: Training, validation and testing results for the rear-view model with no dropouts, learning rate=0.0006, 38 epochs and a learning rate scheduler.

Evaluation matrix	Training	Validation	Testing
MSE	31381,393	379376,847	2191738,391
RMSE	177,148	615,936	1480,452
MAE	134,593	373,632	1148,319
MAPE	-	-	0,173
R^2	0,992	0,897	0,302

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R^2 = coefficient of determination.

The drop from the model validation performance to the testing performance is most likely due to validation done on the augmented data. This drastically escalates the validation results, albeit training the model by making it invariant to image rotations, colour changes or where the cow is facing the image.

Ozkaya (2015) predicted milk yield using rear udder images and reported a better coefficient of determination of 0.66 than the R^2 of 0.3 found in the current study. The differences in the performance can be attributed to the non-standardization of factors such as the milking time, cow's standing posture, camera pixel resolution, illumination conditions, camera location and settings and distance from the cows in this study. The aim of not standardizing these factors was to create a model that is invariant to changes in factors such as how full the udder is with milk, the cow's standing posture, camera resolution, illumination in the environment, and

distance of the cow from the camera. This, however, caused a reduction in the model performance.

Ozkaya's analysis process was more supervised compared to this study. In the current study, the features to extract from the image and the trait weights were unknown. This suggests that some level of supervised learning may be necessary to give better predictions of milk yield from image data with the same data size.

The model proposed in the current study gave a lower prediction accuracy than other models that used linear-type traits to predict milk yield, such as the ones used by Yakubu (2011) with an R^2 of 0.69 when all conformation traits evaluated were considered and Gocheva-ilieva & Yordanova (2022) with an R^2 ranging from 0.93 to 0.95 when all assessed traits were considered. These studies used a wider range of linear-type traits and not just traits seen from the posterior of the cow. Gocheva-ilieva & Yordanova (2022) considered other factors that are not conformation traits, i.e. lameness, the farm and locomotion (Gocheva-ilieva & Yordanova, 2022) which, in turn, improved the performance of the regression. Yakubu (2011) used heart girth, fore right teat length, fore left teat length, rear right teat length, rear left teat length, udder circumference and udder height, and Gocheva-ilieva & Yordanova (2022) used stature, chest width, rump width, rear legs (rear-view), rear leg set (side-view), hock development, bone structure, foot angle, foot depth and udder width to predict milk yield. Therefore, there may be need to incorporate more conformation traits to the study and not just those viewed from the rear-view.

3.3.1. Effects of splitting the test data after augmentation

Another test was done to establish the effect of splitting the test data after augmentation. All the other parameters used in model training and validation and the data split ratios were the same as in the one where the test data was split before augmentation. The results of the effect of splitting the testing data after augmentation are shown in Table 3.8. The model had an RMSE of 842.6 kg, an MAE of 525.1 kg, an MAPE of 0.08 and an R^2 of 0.82 on the test data. For model training, the MSE, RMSE, MAE and R^2 were 71169.3, 266.8 kg, 216.3 kg and 0.98, respectively. The MSE, RMSE, MAE and R^2 were 669992.2, 818.5, 516.4 kg and 0.82 for the model validation, respectively.

As indicated in the figures above, the model's prediction accuracy was exceptionally better compared to the one where the test data was split before augmentation. The better performance was because the model's test cases were not entirely new. Even though data augmentation improves learning by artificially increasing the sample size and training the model to be

invariant to changes such as rotation, colour and the direction the cow is facing on the image, it cannot be used for the testing data as some of the images are not entirely novel for the model.

Table 3.8: Training, validation and testing results for the rear-view model with the test data split after augmentation, no dropouts, learning rate=0.0006, 38 epochs and a learning rate scheduler.

Evaluation matrix	Training	Validation	Testing
MSE	71169,274	669992,215	709950,502
RMSE	266,776	818,531	842,586
MAE	216,272	516,398	525,116
MAPE	-	-	0,077
R ²	0,981	0,819	0,815

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R²= coefficient of determination.

Figure 3.5 is a scatter-plot from the models where the test data was only split after augmentation and Figure 3.6 is for a model where the test set was split before augmentation. The y-axis represents actual milk yield values and is colour coded from light blue to purple. Light blue dots indicate cows with the lowest milk yield, and purple dots indicate cows with the highest milk yield. The x-axis shows the predicted milk yield. The data and parameters were the same for these two figures. The difference in performance was due to data memorization in Figure 3.5.

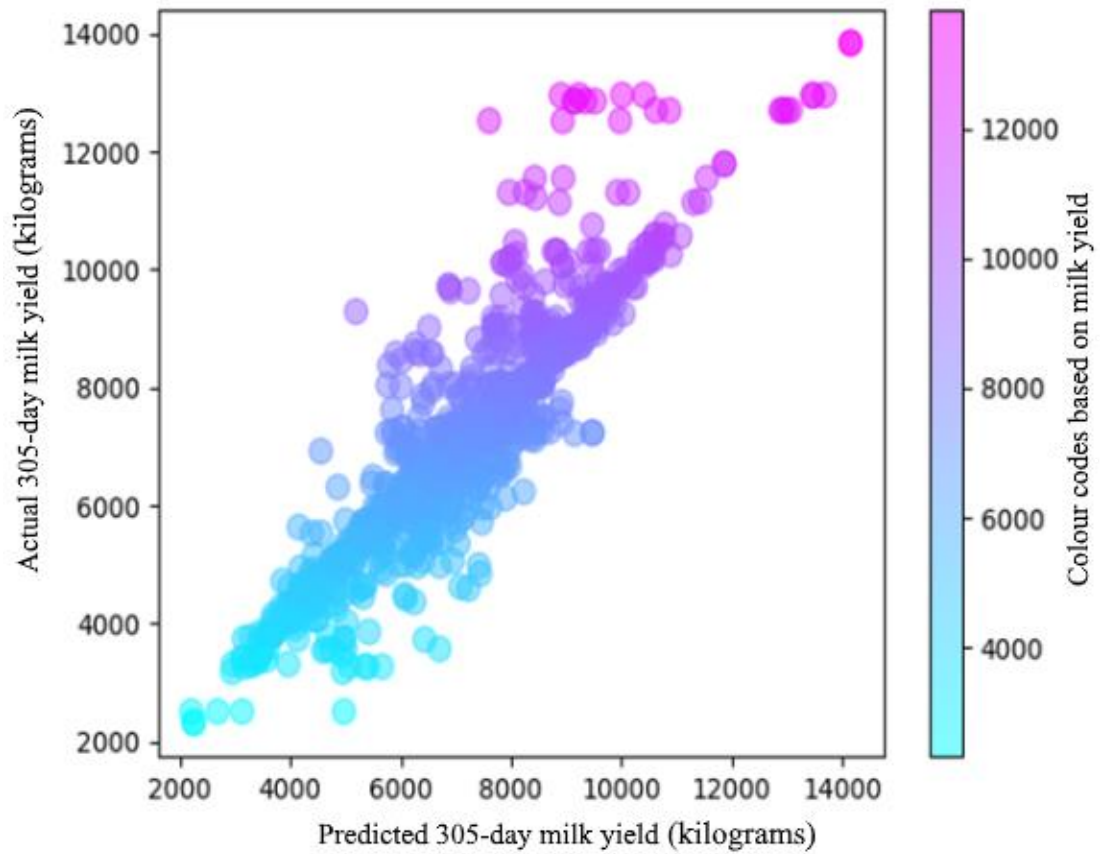


Figure 3.5: Scatter plot of the predicted vs the actual milk yield values from the rear-view model where the test data is split after augmentation.

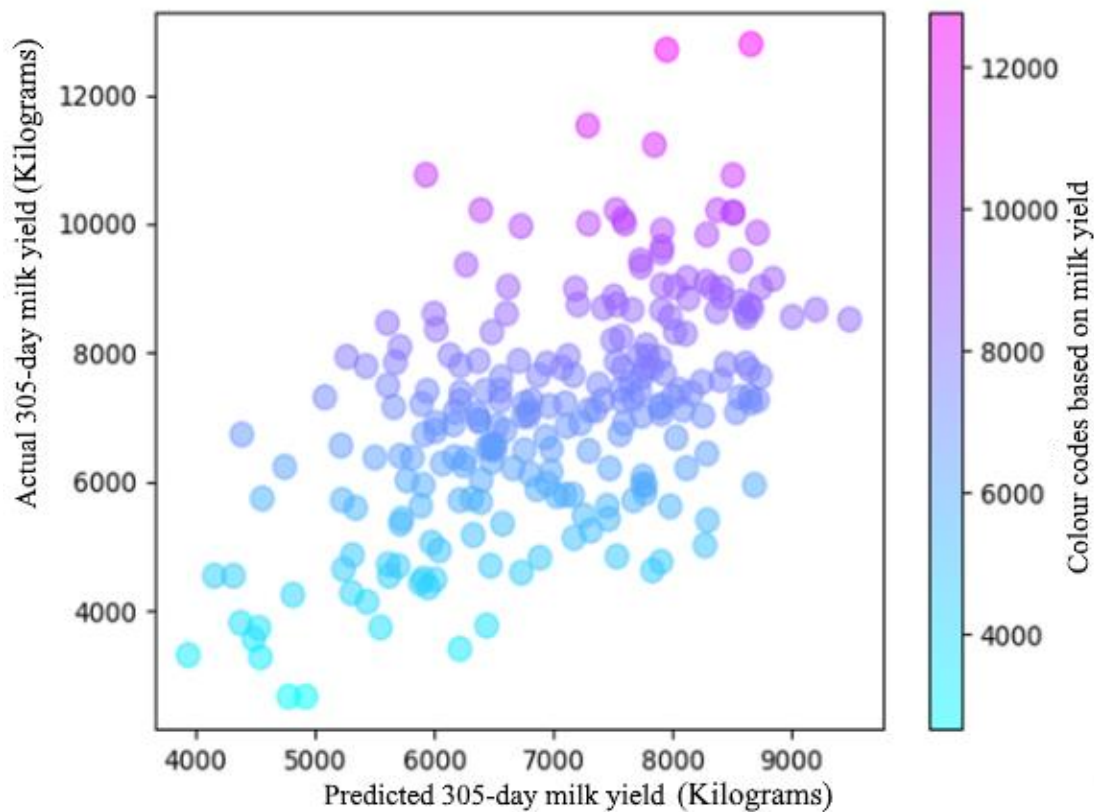


Figure 3.6: Scatter plot of the predicted vs the actual milk yield values from the rear-view model where the test data is split before augmentation.

As seen in figure 3.6, where the data was split before augmentation, there were no predictions beyond ten thousand kilograms. On the memorized data, shown in figure 3.5, predictions were poorer for cows with milk yield greater than ten thousand kilograms per lactation. This is seen through the scattering of the purple dots across the x-axis. Poor predictions for high-producing cows could have been due to a limited number of cows within this range for adequate model training or because there are minimum trait differences between high-producing and very high-producing cows. Other factors such as management, feed availability and quality, and other genetic factors such as feed conversion efficiency could be impelling the cows to produce exceptional amounts of milk (Gross. 2022). This implies that for outstanding milk production, farmers need not only concentrate on selecting genetically superior breeds but also improve on management, quality of feed and its availability.

3.4. Conclusions

The 305-day milk yield of cows is predictable using rear-view images. The predictions are, however, generally weak. The weak predictions could have resulted from factors such as variations in the fullness of the udder with milk, full automation on feature selection, and

differences in the camera's distance from the cow and posture differences, especially regarding the rear leg position. There is need to assess whether lactation milk yield can be predicted using side-view images of milking cows.

3.5. References

- Andrew, S., Georg, H., Marc'Aurelio, R., & Ke, Y. (2013). An empirical study of learning rates in deep neural networks for speech recognition. New York, 6724–6728.
- Campos, R. V., Cobuci, J. A., Kern, E. L., Costa, C. N., & McManus, C. M. (2015). Genetic parameters for linear type traits and milk, fat, and protein production in Holstein cows in Brazil. *Asian-Australasian Journal of Animal Sciences*, 28(4), 476–484. <https://doi.org/10.5713/ajas.14.0288>
- Chollet, François. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. arXiv. Doi:10.48550/arXiv.1610.
- Getu, A., & Misganaw, G. (2015). The Role of Conformational Traits on Dairy Cattle Production and Their Longevities. *OALib*, 02(03), 1–9. <https://doi.org/10.4236/oalib.1101342>
- Gocheva-ilieva, S., & Yordanova, A. (2022). Predicting the 305-Day Milk Yield of Holstein-Friesian Cows Depending on the Conformation Traits and Farm Using Simplified Selective Ensembles. *Mathematics*. <https://doi.org/10.3390/math10081254> Academic
- Harris, R. (2015). Phenotypic correlations between linear type conformation traits, production and fertility in a once-a-day milked dairy cattle herd. Thesis retrieved 5 May, 2021 from the Massey university website: https://mro.massey.ac.nz/bitstream/handle/10179/7430/02_whole.pdf?sequence=2&isAllowed=y.
- International Committee of Animal Recording. (2015). 1 . Conformation Recording of Dairy Cattle. June, 1–42. retrieved 1 August, 2021 from the ICAR website: <https://www.icar.org/wp-content/uploads/2015/08/Conformation-Recording-CR-WG.pdf>
- Janković, D., Trivunović, S., & Samolovac, L. (2020). Characteristics of linear traits of udder and angularity in Holstein- friesian cows and their correlation with milk yield traits. 36(4), 407–416.
- Josef J Gross. (2022). Limiting factors for milk production in dairy cows: perspectives from physiology and nutrition, *Journal of Animal Science*, Volume 100, Issue 3, March 2022, skac044, <https://doi.org/10.1093/jas/skac044>
- Kandel, I., & Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, 6(4), 312–315. <https://doi.org/10.1016/j.icte.2020.04.010>

- Kern, E. L., Cobuci, J. A., Costa, C. N., McManus, C. M., Campos, G. S., & Almeida, T. P. (2015). Genetic association between longevity and linear type traits of holstein cows. *Scientia Agricola*, 72(3), 203–209. <https://doi.org/10.1590/0103-9016-2014-0007>
- Khan, M. S. M. A., & Khan, M. S. M. A. (2016). Genetic and phenotypic correlations between linear type traits and milk yield in Sahiwal cows. *Pakistan Journal of Agricultural Sciences*, 53(2), 483–489. <https://doi.org/10.21162/PAKJAS/16.3369>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 1–15.
- Majurski, M. (2019) ‘Small Data Deep Learning : AI Applied to Domain Datasets’, Nist. Retrieved 21 August 2021 from the National Institute of Standards and technology website: https://www.nist.gov/system/files/documents/2019/08/27/workshopslides-small_data_convnets.pdf
- Ozkaya, S. (2015). Prediction possibility of milk yield from udder measurements using digital image analysis on Holstein cows. *Indian Journal of Animal Research*, 49(3), 388–391. <https://doi.org/10.5958/0976-0555.2015.00050.3>
- Tapki, I. (2013). Genetic and Phenotypic Correlations between Linear Type Traits and Milk Production Yields of Turkish Holstein Dairy Cows By. *Green Journal of Agricultural Sciences*, 3(11), 755–761. <https://doi.org/10.15580/GJAS.2013.11.072913763>
- World weather online. (2022). Humansdorp Climate Weather Averages. Retrieved 5 January 2022 from the world weather inline website: <https://www.worldweatheronline.com/humansdorp-weather-averages/eastern-cape/za.aspx>
- Zindove, T. J., Chimonyo, M., & Nephawe, K. A. (2014). Relationship between linear type and fertility traits in Nguni cows. *Animal*, 9(6), 944–951. <https://doi.org/10.1017/S1751731114003231>

CHAPTER FOUR: Predicting milk yield using a cow's side-view image through computer vision deep learning

Abstract

The objective of this current study was to determine the predictability of cows' 305-day milk yield through computer vision deep learning using cow's side-view images. One thousand two hundred thirty-eight side-view images from seven hundred and forty-three Holstein cows within their first or second parity were used together with their corresponding first lactation milk yield values. The data were split at a ratio of 80:20 into the training and testing data, respectively. The training data was augmented, increasing the training data to 5005 images. These five thousand and five images were again split into training and validation data at the ratio of 80:20, respectively. The training and validation data were used to train and validate the deep learning model, and the test data to check the model's performance on original, unaugmented data. The Xception architecture was the pre-trained Convolutional Neural Network used. From the testing data, the results observed were 2133762.6 kg; 1460.7 kg, 1146.4 kg, 0,18 % and 0,32 for the MSE, RMSE, MAE, MAPE and the R² value, respectively. The weak predictions imputed to the complete automation of the prediction process, especially in image feature selection, differences in cows' standing postures and the camera's distance from the cow. The performance of this side-view images model was compared to that of the model of the rear-view image. No significant difference was found between the two models ($p>0.05$). This was possibly due to correlations among the side-view and the rear-view conformation traits. The conclusions were that predictions made from side-view images when using computer vision deep learning are poor, and the linear-type traits seen from the cow's side, when considered as gestalt, provide the same milk yield information as those seen from the posterior of the cow.

Keywords: 305-day milk yield, deep learning, side-view images, linear-type traits, Xception network

4.1. Introduction

The profitability of a dairy farm lies in good management, and good management entails foresight and sound decision-making in mating, feeding and culling. One of the foresight required in dairy production is the amount of milk a cow will produce in a given time. Not only does this enhance the decision-making process, but it also gives an insight into the animals' health (Liseune et al., 2020).

Milk yield can be predicted using various ways, which include using genomics, early MY records of an animal, pedigree records and linear conformation traits. Most MY predictions done using linear-type traits tend to focus more on rear-view traits such as udder width, udder height and rump width (Yakubu, 2011; Ozkaya, 2015). This gives a perception that when using cow image data, a cow's rear-view image is a better predictor of milk yield than the side-view is. Needless to say that there are side-view traits that have a moderate to high correlation with milk yield, such as angularity (Zink et al., 2014; Campos et al., 2015; Harris, 2015; Khan & Khan, 2016). There are few, if any, pieces of research comparing the predictability of milk yield from side-view traits vs from the rear-view traits. This chapter aims to determine the predictability of milk yield from the side-view images and compare the results to milk yield prediction using the rear-view.

In as much as some research only review the predictability of milk yield from the rear-view traits, it may be inaccurate to assume that the rear-view is more critical in predicting MY than the side-view based on a priori knowledge because the order of importance of these conformation traits in explaining milk yield variability varies widely in the literature. Therefore, there is a need to do a practical evaluation comparing the performance of the side-view images to that of the rear-view images in predicting milk yield. This chapter evaluates milk yield prediction from only the side-view images through computer vision deep learning and compares the results with those from the rear-view images.

4.2. Hypothesis

The hypothesis tested in this study was:

1. Milk yield prediction is the same when using either the rear-view or the side-view images of cows.

4.3. Materials and Methods

4.3.1. Description of the study sites

The study sites were described in detail in chapter 3.2.1.

4.3.2. The study variables

The dependent variable in this study was milk yield, and the independent variable was the cows' side-view images. The control variables were the number of milkings, feeding system and image dimensions.

4.3.3. The data

Unlike the description of data in chapter 3.2.3, which used the rear-view images of cows, the data required here consisted of cows' side-view images and the cows' corresponding 305-day milk yield values from the first lactation. From the side-view, the desired features were, but not limited to, angularity and udder depth. Since cow visual images were used and no feature extraction was done, some features not named above may have impacted the overall milk yield prediction. The remaining information about the data and data collection is exactly as described in chapter 3.2.3.

4.3.4. Data pre-processing

Image pre-processing for the side-view images and the creation of CSV files is as described in chapter 3.2.4.

4.3.5. Data analysis

Data analysis was split into hyperparameter optimization and the actual tests using the optimal parameters. This study was carried out using the Xception deep learning network/model (Chollet, 2017). The network was implemented using the TensorFlow framework. Due to insufficient GPU space, the models were trained on a Central Processing Unit (CPU) allocator. A CSV file with a milk yield column and the side-view image column was used to call the images and their corresponding MY values to the model.

The Mean Square Error (MSE), mean absolute error (MAE), and the coefficient of determination (R^2) values were used to evaluate and update trait weights. The root mean square error (RMSE), MAE, Mean Absolute Percentage Error (MAPE), and R^2 values were used to report the research findings on the test data. The equations for the MSE, RMSE, MAE MAPE and the R^2 are given in chapter 3.2.5.

A hypothesis test was then done to check whether there was any significant difference in performance between the side-view model and the rear-view model. To enable the t-test, the observed and predicted values were recorded from each model. The absolute error terms between the observed and the predicted milk yield values were then calculated for each record. A two-tailed, paired t-test was used to compare the two models based on the error values of each. The t-test was two-tailed because we were interested in determining differences in any direction. Either the side-view model is better than the rear-view model or vice versa. Because the testing sample contained the same cows in both models, a paired t-test was used. Below is the formula for the t-test used:

$$t = \frac{\bar{X}_D}{S_D / \sqrt{n}}$$

Where: \bar{X}_D = the average of the differences between the models' error terms (i.e. absolute difference between observed and predicted MY values for each model)

and S_D = the standard deviation of the differences between the models' error terms.

Hypothesis (H0):

There are no significant differences ($P > 0.05$) in milk yield prediction values based on side-view images of cows and prediction based on rear view images.

4.3.5.1. Hyperparameter optimisation and the optimum parameters used

The parameters examined were the best learning rate, whether or not to use a learning rate decay, whether or not to use dropouts, and the optimum number of epochs. Details on what parameter optimization is and the function of each parameter are given in chapter 3.2.5.1.

Because of lower CPU memory, the batch size remained unchanged for all the experiments in this research. A batch size of 32 was used. Learning rates between 0.0004 and 0.001 were tried for the side-view analysis model. Based on all the evaluation matrices, the optimum learning rate was 0.0005. The LR comparison results were done at 35 epochs with no LR scheduler or dropout. Table 4.1 shows a comprehensive report of the comparison between various learning rates. Below Table 4.1 is Figure 4.1, plotted based on the validation MAE from Table 4.1, clearly showing the LR with the lowest MAE.

Table 4.1: Side-view learning rate comparison results at 35 epochs, with LR scheduler and no dropout.

LR	Training				Validation			
	MSE	RMSE	MAE	R ²	MSE	RMSE	MAE	R ²
0,0004	328567,835	573,208	532,387	0,913	732602,532	855,922	676,440	0,801
0,0005	44423,388	210,769	168,757	0,988	372958,750	610,703	376,715	0,898
0,0006	47703,177	218,411	170,476	0,987	558473,383	747,311	454,137	0,848
0,0007	54702,931	233,887	181,647	0,986	477490,574	691,007	425,082	0,870
0,0008	112779,581	335,827	289,704	0,970	652901,253	808,023	533,260	0,822
0,0009	279367,686	528,552	474,509	0,926	831673,045	911,961	698,581	0,774
0,001	74678,319	273,273	212,597	0,980	587384,168	766,410	491,085	0,840

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R²= coefficient of determination.

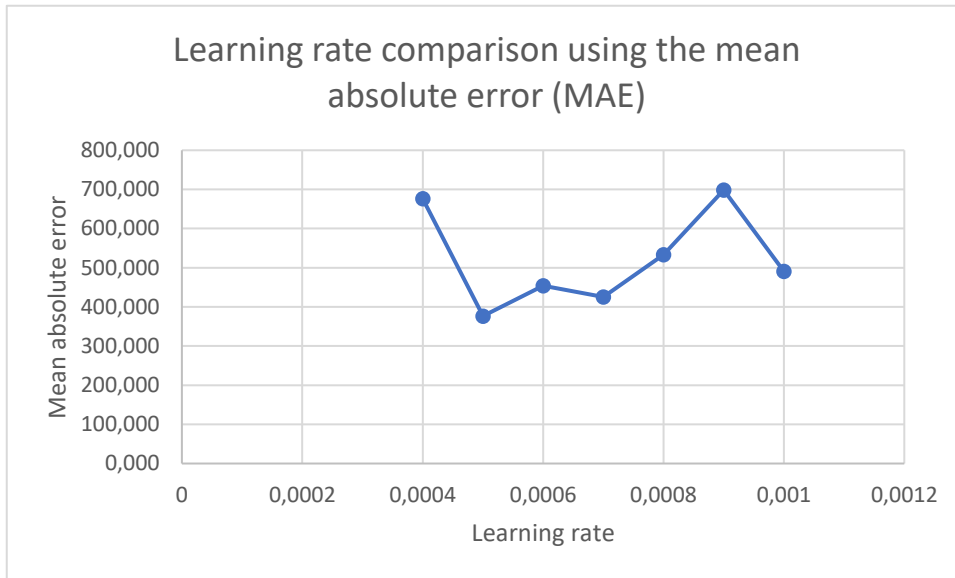


Figure 4.1: Side-view learning rate comparison results at 38 epochs, with LR scheduler and no dropout.

In a practical comparison reported in Table 4.2, based on all evaluation metrics, better results were observed when a manual LR scheduler was absent. Tests to examine whether or not to use an LR scheduler were done using the optimum Learning rate of 0.0005, no dropouts, and 35 epochs.

Table 4.2: Comparison between a side-view model with a learning rate scheduler and one with no LR scheduler.

Evaluation matrix	LR decay present	LR decay absent
MSE	472376,944	391588,271
RMSE	687,297	625,770
MAE	464,541	364,065
R ²	0,871	0,893

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R²= coefficient of determination.

The Adam optimizer was used for all models in this research because it is currently the best-performing optimizer in image and language processing. The model was run at fifty epochs to establish the best number of epochs. The purpose was to observe the epoch range at which the validation curve stabilizes. About thirty-five epochs were found to be optimum for the side-

view. Figure 4.2 shows the learning curves for training and validation based on the MAEs. The test was done with the LR of 0.0005 and with a learning rate schedule present.

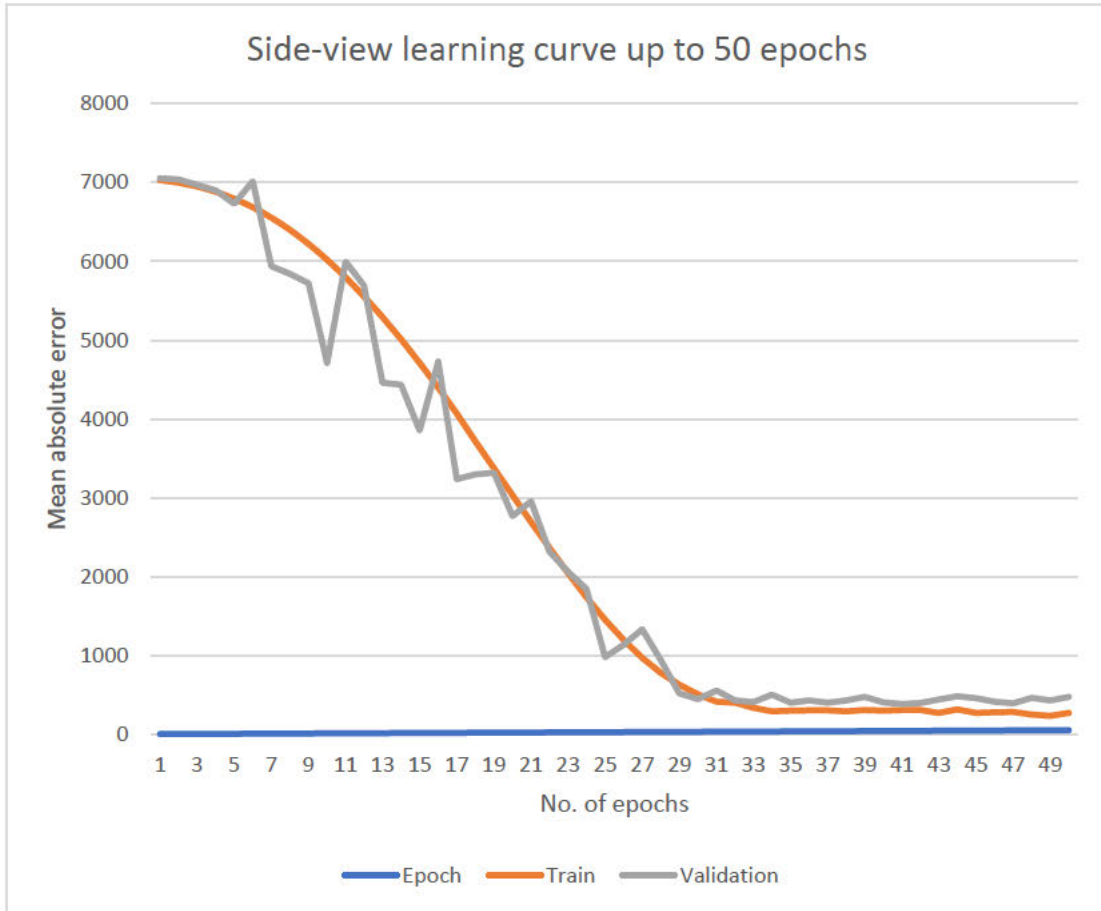


Figure 4.2: Side-view learning curves for the training and validation.

A model with a dropout rate of 0.4 was compared to those with no dropout. From all evaluation metrics, i.e. MSE, RMSE, MAE and R^2 values, better results were seen in models with no dropouts. The tests were done with an LR of 0.0005 at 35 epochs. Table 4.3 compares a model with dropouts and one without dropouts.

Table 4.3: effects of dropouts (dropout rate=0.4) on model learning.

Evaluation matrix	Dropout present	Dropout absent
MSE	442269,424	372958,750
RMSE	665,033	610,703

MAE	388,655	376,715
R ²	0,880	0,898

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R²= coefficient of determination.

In summary, the optimiser and batch size used was the same for all models and was determined based on a priori knowledge. The Adam optimizer and a batch size of 32 were used for the final model. Based on experimental evidence for the side-view model, dropouts were not used, an LR of 0.0005 was used with no LR scheduler, and 35 epochs were used. Table 4.4 summarises the hyperparameters used for the final rear-view model.

Table 4.4: Optimum parameter used for the final side-view analyses.

Hyperparameter	Side-view analysis
Optimiser	Adam
Initial learning rate	0.0005
Learning rate scheduler	No
Batch size	32
The final epochs number	35
Drop out	No

4.3.6. Description of the data

The description of the data in terms of the sample size, percentages of data used for training, validation and testing, the frequency distribution for the training and testing datasets, average milk yield, maximum MY, minimum MY, and standard deviation of the data is precise as described in chapter 3.2.6. The side-view images and their corresponding 305-day milk yield values were considered.

4.4. Results and Discussion

The final model for the prediction of milk yield from the cows' side-view images was then run on testing data. The results show that deep learning can be used to predict milk yield using side-view images. The model accuracy, however, dropped by about threefold on the testing data compared to the validation data. For the validation, the values were 391588.3; 625,8 kg; 364,1 kg and 0,89 for the MSE, RMSE, MAE and the R² value, respectively. The results from the testing data were 2133762,6; 1460,7 kg; 1146,4 kg; 0,18 % and 0,32 for the MSE, RMSE; MAE, MAPE and the R² value, respectively. Table 4.5 shows the side-view model's training, validation, and testing results at the established optimum parameters.

The drop in performance from validation to testing resulted from the validation being done on data containing augmented images and the model testing done on real unaugmented images. The accuracy of milk yield prediction from traits viewed from the side of a cow discovered in this research was generally low. As much as various research points out the correlation between some traits seen from the side of a cow, such as angularity and stature (Zink et al., 2014; Campos et al., 2015; Harris, 2015; Khan & Khan, 2016), there are few pieces of research, if any, addressing the predictability of milk yield based only on these traits. The low prediction accuracy can be imputed to either the absence of traits that adequately explain milk yield variability or the use of computer vision deep learning and the training data size.

Table 4.5: Training, validation, and testing results for the side-view model with no dropouts, a learning rate of 0.0006, 38 epochs and a learning rate scheduler.

Evaluation matrix	Training	Validation	Testing
MSE	21330,468	391588,271	2133762,556
RMSE	146,050	625,770	1460,74
MAE	111,144	364,065	1146,356
MAPE			0,176
R ²	0,994	0,893	0,320

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error, MAPE= Mean absolute percentage error and R²= coefficient of determination.

As much as milk yield can be predicted using side-view images through deep learning, farmers may need to use this method together with other prediction methods, such as the use of early milk yield records, for better prediction results. There is also a need to investigate whether the low prediction accuracy was due to the training method or the unavailability of conformation traits to adequately explain variability in milk yield. A comparison between the side-view and rear-view models' prediction accuracy could answer the question of whether the low prediction accuracy was due to the inadequacy of traits to explain the variability in milk yield or it was because of the prediction method used.

4.4.1.Side-view model vs rear-view model

The rear-view analysis discussed in the previous chapter used the same prediction method to evaluate the prediction of MY. A hypothesis test comparing the side-view model's performance to the rear-view model can indicate whether the low prediction accuracy found in this study is a conformation-traits problem or a prediction method problem. If the side-view and rear-view models are comparable or the side-view model gives higher MY predictions, then the poor performance of the side-view prediction model can be attributed to the use of deep learning rather than the inability of the side-view conformation traits to predict milk yield.

A hypothesis test was done to compare the performance of the side-view model to the rear-view and check if there is any significant difference. The table below is a summary of the two-tailed paired t-test done. As shown in Table 4.6 the means were 1146.4 kg for the side-view data and 1148.3 kg for the rear-view data. The variances were 823121.6 and 876889.9 for the side-view and rear-view models, respectively. There was no significant difference in performance between the side-view and the rear-view model ($p > 0.05$). This means the weight of side-view traits as gestalt in explaining variability in MY is the same as the rear-view traits weight. In other words, the side-view image is as good/poor as the rear-view in explaining variability in milk yield. The Pearson correlation between the two prediction models was 0.66 and the P-value was 0.97.

Table 4.6: T-test: Two-tailed, paired t-test between the side-view model and the rear-view model.

	Side view	rearview
Mean	1146,388	1148,338
Variance	823121,552	876889,852
Observations	237	237
Pearson correlation	0,659	
Hypothesized mean difference	0	
df	236	
t Stat	-0,039	
P(T<=t) two-tail	0,969	
t Critical two-tail	1,970	
Results	NS	

df: degrees of freedom. *=Significantly different at ($P < 0.05$). NS= Not significant ($P > 0.05$)

4.4.2. Effects of splitting the test data after augmentation

Apart from the data split sequence, the other parameters used in this model training and validation were the same as in the test where the test data was split before augmentation. As shown in Table 4.7, the model had an RMSE of 778.6 kg, an MAE of 492.6 kg, a MAPE of 0.07 and an R^2 of 0.84. This shows that the model's test cases were not entirely new. Hence the model's prediction accuracy was better than the one where data splitting was done before augmentation.

Table 4.7: Training, validation and testing results for the side-view model with the test data split after augmentation, no dropouts, no learning rate scheduler, initial learning rate= 0.0005 and 35 epochs.

Evaluation matrix	Training	Validation	Testing
MSE	54455,293	768111,394	606160,730
RMSE	233,357	876,420	778,563
MAE	181,540	528,139	492,585
MAPE			0,072
R-Square	0,985	0,793	0,842

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error, MAPE= Mean absolute percentage error and R^2 = coefficient of determination.

Figure 4.3 is a scatter plot from the model where the test data was only split after augmentation, and Figure 4.4 is for a model where the test set was split before augmentation. The y-axis represents actual milk yield values, and the x-axis shows the predicted milk yield. The colour codes are for the actual milk yield values. Light blue dots are for cows with the lowest milk yield, and purple dots are for cows with the highest milk yield.

The dots in Figure 4.3, representing each cow, was more linear than in Figure 4.4. This means predictions shown in Figure 4.3 were more accurate than those in Figure 4.4. As shown in Figure 4.3, the purple dots, representing higher milk yield cows, were more scattered on the x-axis, showing reduced prediction accuracy for cows in the purple range.

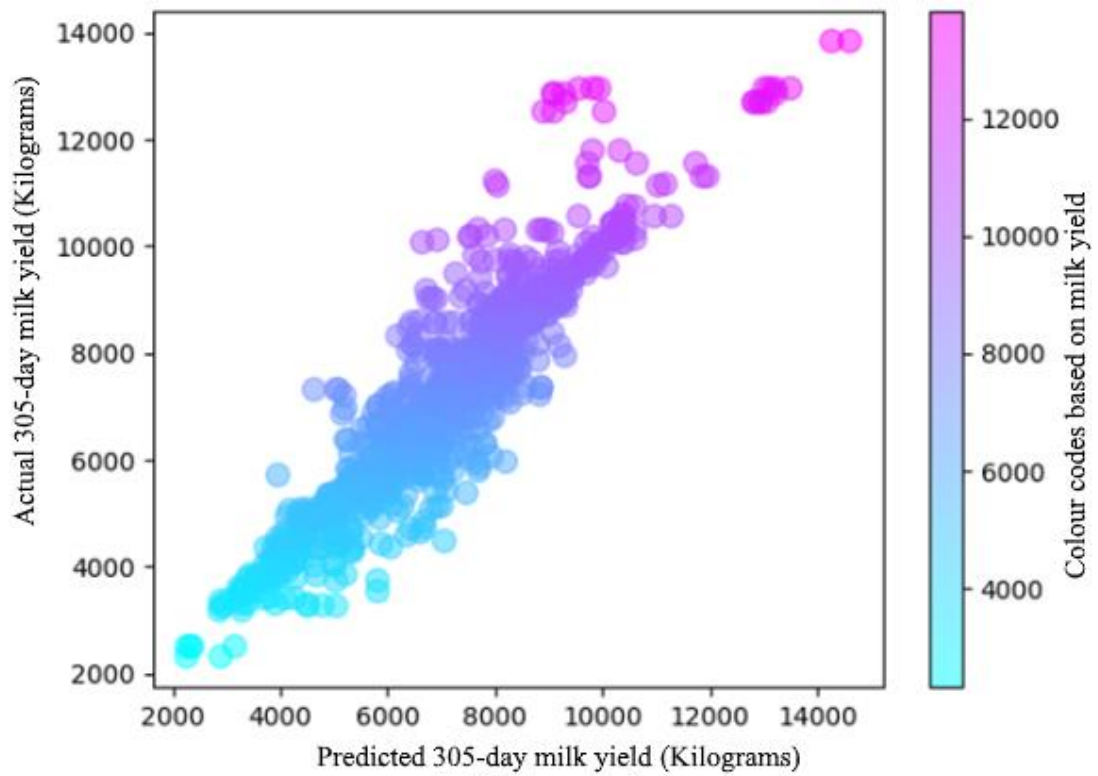


Figure 4.3: Scatter plot of the predicted vs the actual milk yield values from the side-view model where the test data is split after augmentation.

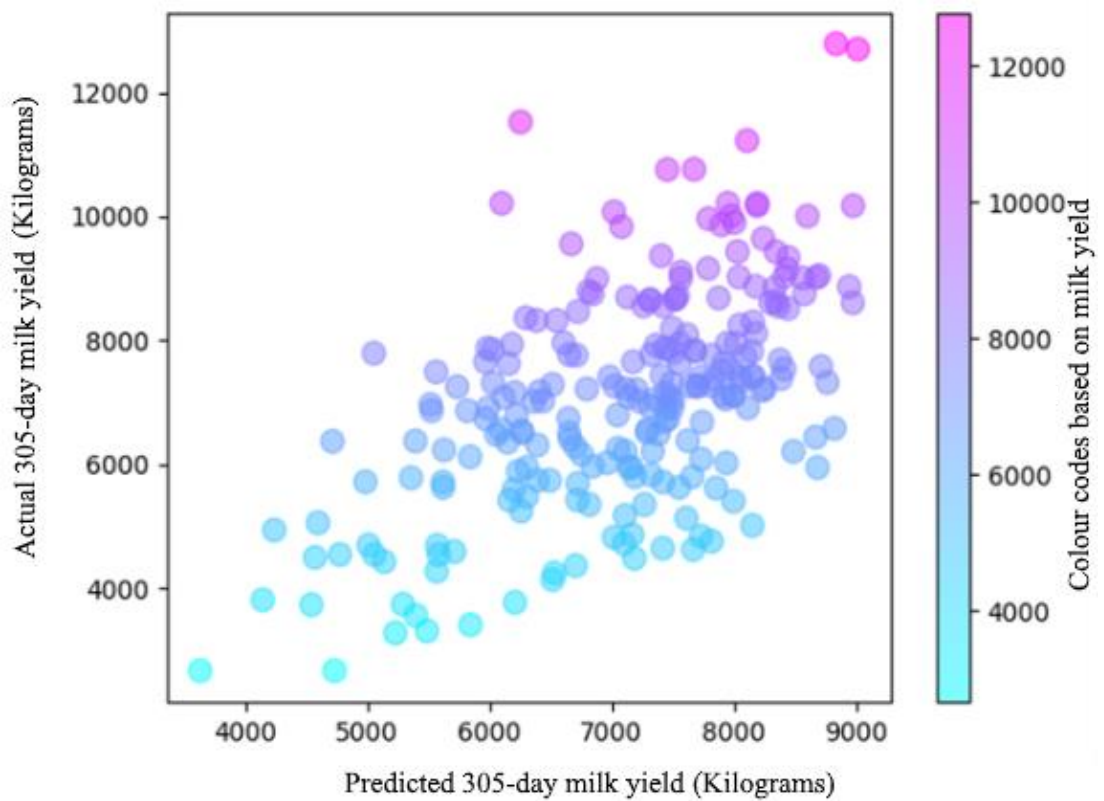


Figure 4.4: Scatter plot of the predicted vs the actual milk yield values from the side-view model where the test data is split before augmentation.

Since the data and parameters were the same for these two figures, the difference in performance between the two models was imputed to data memorization on the model where the test data was only split after augmentation (Shorten & Khoshgoftaar, 2019). For the memorized data, predictions were poor for cows with milk yield greater than ten thousand kilograms. This is either because of inadequate data within this range for model training or because there is no difference in appearance between cows within this MY range. Differences in milk yield could result from other factors such as feed availability and quality, management, and other genetic factors such as feed conversion efficiency (Gross. 2022).

4.5. Conclusions

The 305-day milk yield is predictable using cows' side-view images. The predictability, however, is generally weak. The weak predictions are probably a result of factors such as the complete automation of the prediction process, especially in image feature selection, differences in cows' standing postures and differences in the camera's distance from the cow. The reason for not standardizing camera distance from the cow and cow postures was to create a model that is indifferent to distance or posture. However, this came at a considerable performance cost. Based on the hypothesis test, the side-view image is as good/poor as the rear-view in explaining variability in milk yield. Further investigations should consider whether combining both the rear- and side-views increases the strength and accuracy of predicting 305-day milk yield.

4.6. References

- Campos, R. V., Cobuci, J. A., Kern, E. L., Costa, C. N., & McManus, C. M. (2015). Genetic parameters for linear type traits and milk, fat, and protein production in Holstein cows in Brazil. *Asian-Australasian Journal of Animal Sciences*, 28(4), 476–484. Doi: 10.5713/ajas.14.0288
- Chollet, François. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. arXiv. Doi:10.48550/arXiv.1610.02357.
- Harris, R. (2015). Phenotypic correlations between linear type conformation traits, production and fertility in a once-a-day milked dairy cattle herd. Thesis retrieved 5 May, 2021 from the Massey university website: https://mro.massey.ac.nz/bitstream/handle/10179/7430/02_whole.pdf?sequence=2&isAllowed=y.
- Josef J Gross. (2022) Limiting factors for milk production in dairy cows: perspectives from physiology and nutrition, *Journal of Animal Science*, Volume 100, Issue 3, March 2022, skac044, <https://doi.org/10.1093/jas/skac044>
- Khan, M. S. M. A., & Khan, M. S. M. A. (2016). Genetic and phenotypic correlations between linear type traits and milk yield in Sahiwal cows. *Pakistan Journal of Agricultural Sciences*, 53(2), 483–489. Doi:10.21162/PAKJAS/16.3369
- Liseune, A., Salamone, M., Van den Poel, D., Van Ranst, B., & Hostens, M. (2020). Leveraging latent representations for milk yield prediction and interpolation using deep learning. *Computers and Electronics in Agriculture*, 175(July), 105600. Doi:10.1016/j.compag.2020.105600
- Ozkaya, S. (2015). Prediction possibility of milk yield from udder measurements using digital image analysis on Holstein cows. *Indian Journal of Animal Research*, 49(3), 388–391. Doi:10.5958/0976-0555.2015.00050.3
- Shorten, C. and Khoshgoftaar, T. M. (2019) ‘A survey on Image Data Augmentation for Deep Learning’, *Journal of Big Data*, 6(1). doi: 10.1186/s40537-019-0197-0.
- Yakubu, A. (2011). Path analysis of conformation traits and milk yield of Bunaji cows in smallholder ’ s herds in Nigeria. *Agricultura Tropica et Subtropica*, 44(3), 152–157.
- Zink, V., Zavadilová, L., Lassen, J., Štípková, M., Vacek, M., & Štolc, L. (2014). Analyses of genetic relationships between linear type traits, fat-to-protein ratio, milk production traits, and somatic cell count in first-parity Czech Holstein cows. *Czech Journal of Animal Science*, 59(12), 539–547. Doi:10.17221/7793-cjas

CHAPTER FIVE: Predicting milk yield using a combination of cows' side-view and rear-view images through computer vision deep learning

Abstract

The predictability of milk yield based on a combination of cows' side and rear-view images through computer vision deep learning was investigated in this study. The data size of 1238 image pairs (the side-view images and the rear-view images) from 743 Holstein first or second parity cows and their respective first lactation milk yield values were used to train a deep learning model. Each side-view image was concatenated with its corresponding rear-view image to make the combined-view image. The 1238 images were split into the training and testing images at the ratio of 80:20, respectively. Only the training images were augmented with four different augmentations to increase the training samples. The training sample was then split again into training and validation images at the ratio of 80:20, respectively. The Xception architecture (a pre-trained Convolutional Neural Network) was then used to create a milk yield prediction model.

The testing results were 1401.2, 1112.9, 0.17 and 0.38 for the RMSE, MAE, MAPE and the R^2 value, respectively. A t-test for paired observations was also done to compare the performance of this combined view to models which contained only the rear-view images and only the side-view images. In all the comparisons, there was no significant difference in performance between the model containing the combined-view images and those containing either the rear-view images only or the side-view images only ($p>0.05$). This was attributed to correlations among the linear-type traits seen from the side of a cow and those seen from the posterior of a cow. The conclusion was that the predictability of milk yield from a combination of a cow's side-view and the rear-view image is poor using computer vision deep learning.

Keywords: Computer vision, deep learning, Xception architecture, 305-day milk yield and Linear conformation traits.

5.1. Introduction

It is the desire of every dairy farmer to predetermine the amount of milk yield each cow in the herd will produce. This equips the farmer with better decisions on culling, feeding, mating and disease detection, making dairy management more efficient. Milk yield prediction is made through various methods, such as using genomic selection, progeny or pedigree records, individual's own records and linear-type traits. Some methods are known to have consistently

high prediction accuracy. For instance, genomic prediction is generally known to give a high prediction accuracy for milk yield (Ding et al., 2013; Zhang et al., 2022). However, genomic prediction requires expert knowledge, which is often not readily available on the farm. Most farmers often resort to using earlier MY records and linear-type traits to predict milk yield. Narrowing down to the use of linear-type traits, trait measurement and weighting is encumbering, time-consuming, and highly dependent on literature used for correlations and heritability. Therefore, this research evaluates the effect of automating the prediction of milk yield using linear-type traits.

In Chapter three and four, individual sides of a cow were used, the rear-view only and the side-view only. Predictions from individual image views were generally inaccurate. The question is, does the accuracy of prediction increase when all the image views are combined and considered in making MY predictions? Based on Yakubu (2011) and Ozkaya (2015), better milk yield predictions were observed when all conformation traits being analysed were considered. Gocheva-ilieva & Yordanova (2022) determined that only three of the ten linear conformation traits analysed influenced the milk yield prediction. Since the literature shows a lack of consensus in the results involving the use of multiple predictor models to predict milk yield, there is a need to practically evaluate whether combining the image views will improve the prediction performance. This chapter aims to determine the predictability of milk yield from a combination of the side-view images and the rear-view, together called the combined-view, through computer vision DL and compare the performance to that of the individual image views.

5.2. Hypotheses

The following hypotheses were tested at 5 % significance level:

1. milk yield prediction is the same when using either the combined-view or the rear-view images of cows.
2. Milk yield prediction is the same when using either the combined-view or the side-view images of cows.

5.3. Materials and Methods

5.3.1. The study sites

The study sites have been described in detail in chapter 3.2.1.

5.3.2. The study variables

The independent variables in this study were a cows' side-view and rear-view images. The dependent variable was milk yield. Control variables were the number of milkings, the feeding system and image dimensions.

5.3.3. The data

The data consisted of cow side-view and rearview images and the cows' corresponding 305-day milk yield values from the first lactation. The desired rearview features were rump width and rear udder linear-type traits: rear udder depth, height, and width. From the side view, the desired features were stature, angularity and udder depth. It is important to note that some features not named above may have impacted the overall milk yield prediction since no manual feature extraction was done.

The remaining information about the number of milking, milking times, and the sample size and data collection is exactly as described in chapter 3.2.3.

5.3.4. Data pre-processing

Except for the CSV file creation section, the data pre-processing process was as described in chapter 3.2.4.

5.3.4.1. Selection of pictures and CSV files creation

The edited images from all four farms were put into one folder. The names for each image were then extracted and put on a CSV spreadsheet as they appeared on the image. The side view image of a cow corresponded with the rearview image for the same cow. This means that there were two columns, the side view column and the rearview column. Annexed to these two columns was the 305-day milk yield column, to make the third column.

At all the farms, the cows were reared and milked in groups. Filtering out undesired cows at the point of image capturing was more difficult since the animals would be moving. Filtering out of images was, therefore, done after data collection. Some of the cows had no 305-day milk yield records, some were much older, even though they were within the desired parity or lactation number, and some were out of parity or lactation number range. For the reasons above, 1048 images were filtered out (524 pairs), leaving out 2476 usable images (1238 pairs). The CSV file was then updated only to contain the names of the pictures used for the analysis and their corresponding 305-day milk yield values.

Twenty percent of the data were randomly extracted and pasted onto a separate CSV file from the initial CSV file to make the testing data. The remaining 80 % was then used to make the CSV file for the TV data. The TV images were augmented using the training and validation CSV information. The augmented file names and their corresponding milk yields were then annexed to the rows on the TV CSV file. The TV CSV now contained both the TV augmented data and TV unaugmented data. It is important to note that the validation data was a randomly selected 20 % of the TV sample after augmentation. Random splitting of the validation data from the training data was put on the model command; hence the CSV file did not require any splitting to separate the training data and the validation data. All the split ratios used in this research were based on a recommendation by (Majurski, 2019).

5.3.5.Data analysis

Data analysis had two stages, hyperparameter optimisation and the actual tests using the optimal parameters. The architecture used for all the analyses was the Xception deep learning architecture. The models were trained on a Central Processing Unit (CPU) allocator.

All three columns from the CSV files created (chapter 5.2.4.1.) were utilized to call the images and their corresponding milk yield values. The side-view and the rear-view CSV columns were used to combine the images into one image array and the first column contained the milk yield values. The image dimensions were standardized at 299x299 for the architecture. Having two images would mean swashing one dimension of both images to fit into the square, as shown in figure 5.1. This is undesirable because it distorts the proportions of the images used for the other image views.

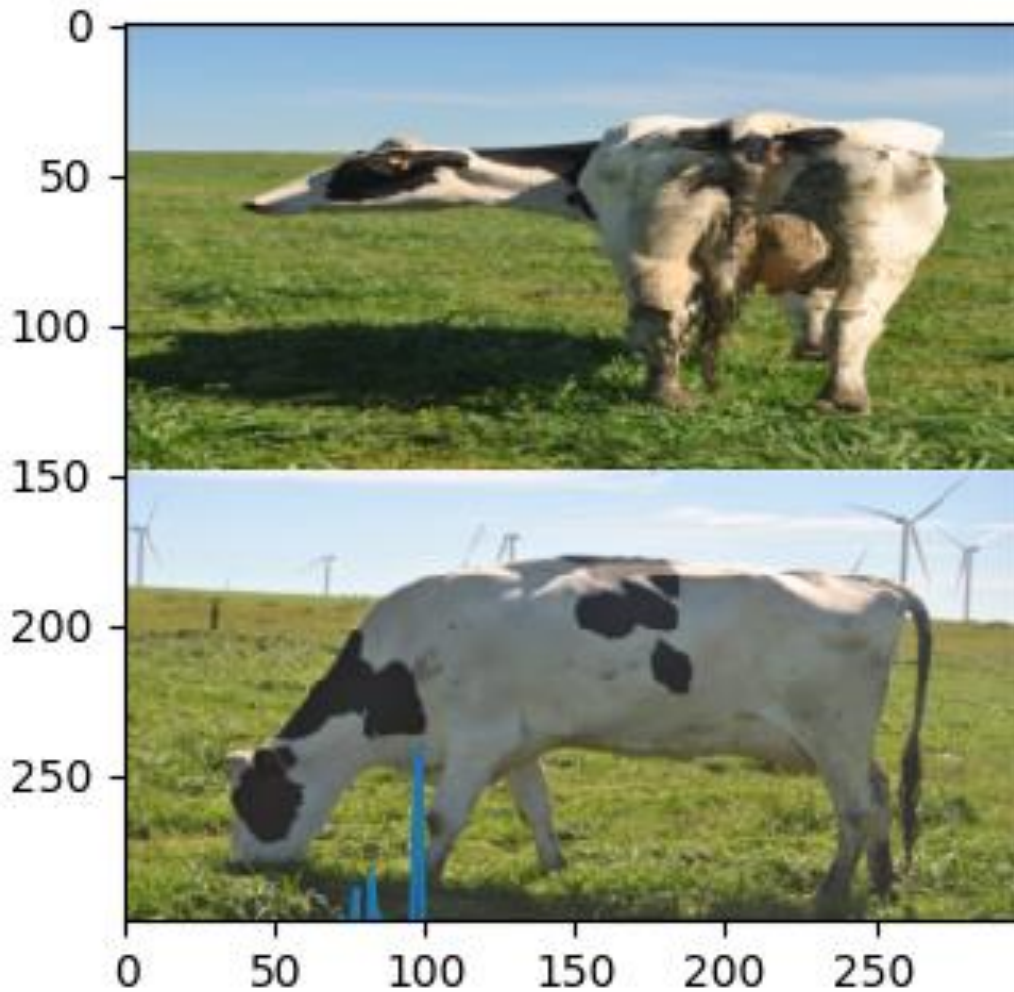


Figure 5.1: Combined-view image with disproportionate side and rear view images.

To solve this problem, side-view images were reduced in dimension to 150x150, and the rear-view images were reduced to 149x149. Two blank images were then created, one for the side-view and the other for the rear-view. The side-view blank image had the dimensions 149x150, and the rear-view blank image had the dimensions 150x149. On top of the final (299x299) image was the side-view blank image, then the side-view image next to it. At the bottom of the final image was the rear-view image on the left and the rear-view blank image on the right. The result was an image with the required dimensions but no stretched images. An example of the output image is shown in Figure 5.2.

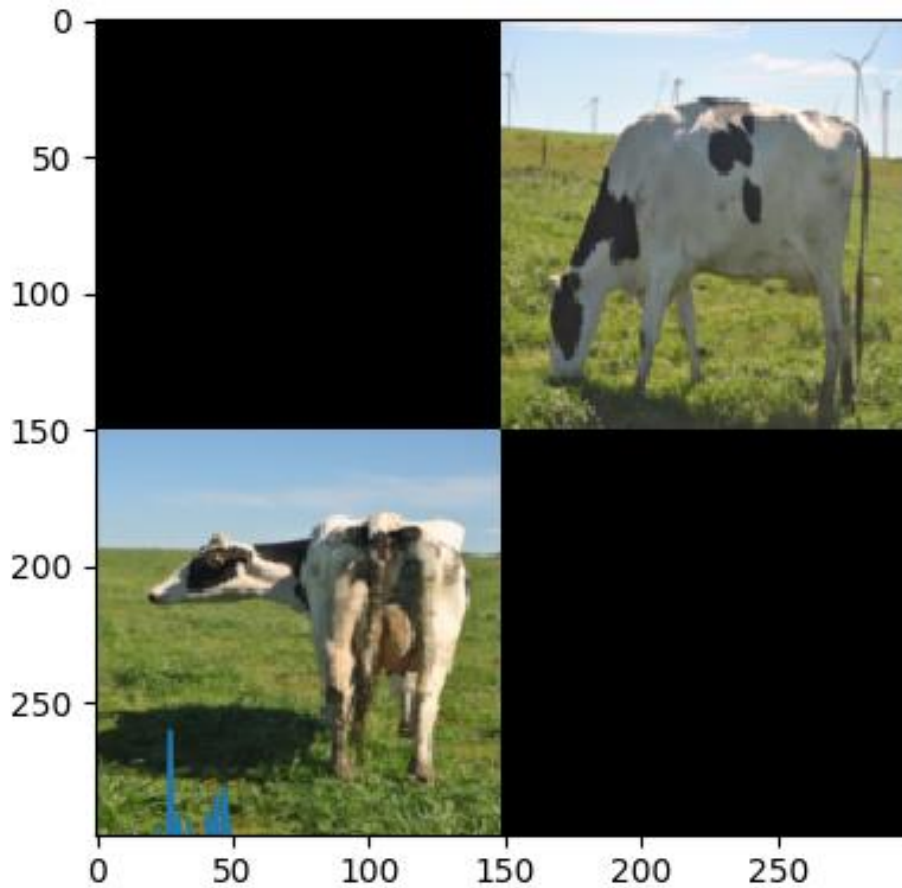


Figure 5.2: Combined images with the same image proportion for the side-view and rear-view

The MSE, MAE, and R^2 values were used to evaluate and update trait weights. The RMSE, MAE, MAPE and R^2 values were used to report the research findings on the test data. The equations for these evaluation matrices were given in chapter 3.2.5.

Two hypothesis tests were then done to determine whether there was any significant difference in performance between the combined-view model and the rear and side view models. A two-tailed, paired t-test was used to compare the two models based on the error values of each model. Below is the formula for the t-test used:

$$t = \frac{\bar{X}_D}{S_D / \sqrt{n}}$$

Where: \bar{X}_D = the average of the differences between the models' error terms (ie absolute difference between observed and predicted MY values for each model)

and S_D = the standard deviation of the differences between the models' error terms.

5.3.5.1. Hyperparameter optimisation and the optimum parameters used

The following hyperparameters were examined, the learning rate to use, whether or not to use a learning rate scheduler, whether or not to use dropouts and the optimum number of epochs. Details on what hyperparameter optimization is and the function of each parameter are provided in chapter 3.2.5.1.

A batch size of 32 was used because of low CPU memory. Model performance for learning rates between 0.0004 and 0.0016 was examined. Based on all the evaluation metrics on the validation data, the optimum learning rate was 0.0006. The LR comparisons were made at 50 epochs with no LR scheduler or dropout. Table 5.1 shows a report on the comparison between the learning rates examined. Below Table 5.1 is Figure 5.3, plotted based on the validation MAE from Table 5.1.

Table 5.1: Combined-view learning rate comparison results at 50 epochs, with no LR scheduler and no dropout.

LR	Training				Validation			
	MSE	RMSE	MAE	R-Square	MSE	RMSE	MAE	R-Square
0,0004	2465047,158	1570,047	1524,028	0,351	2543692,048	1594,896	1524,539	0,308
0,0005	39820,697	199,551	145,070	0,990	290560,379	539,037	359,856	0,921
0,0006	18131,985	134,655	109,472	0,995	245998,056	495,982	312,904	0,933
0,0007	333166,667	577,206	539,474	0,912	630880,041	794,280	636,485	0,828
0,0008	297024,144	544,999	477,194	0,922	733400,403	856,388	666,453	0,800
0,0009	77667,814	278,690	220,377	0,980	340825,937	583,803	403,853	0,907
0,001	119210,419	345,269	274,193	0,969	513174,625	716,3625	501,977	0,860
0,0011	34540,810	185,852	153,738	0,991	332775,333	576,867	355,558	0,909
0,0012	228395,913	477,908	442,024	0,940	666630,495	816,474	640,360	0,819
0,0014	455399,196	674,833	598,952	0,880	838157,691	915,510	730,796	0,772
0,0016	2363298,804	1537,302	1489,245	0,373	2253692,298	1501,230	1302,583	0,303

LR= Learning rate, MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R^2 = coefficient of determination.

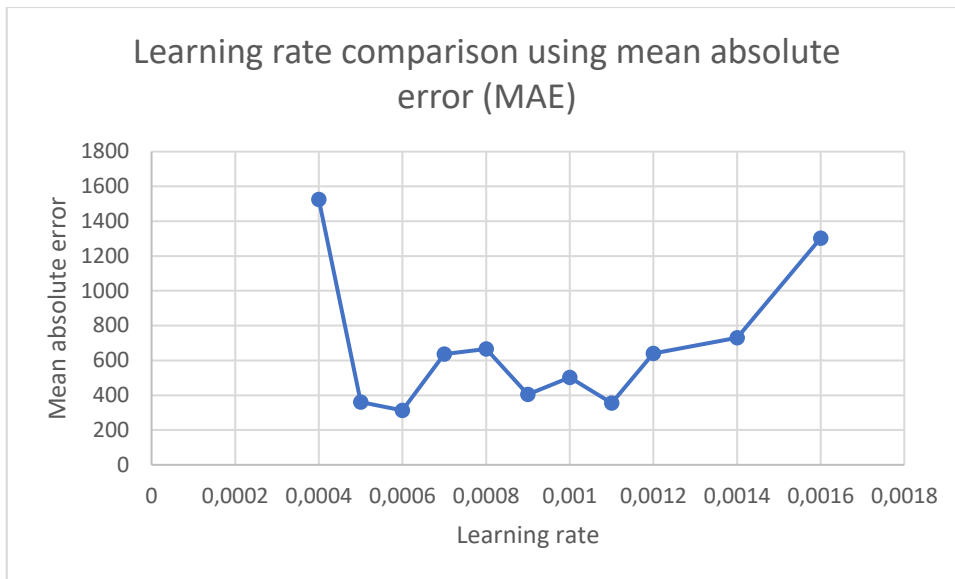


Figure 5.3: Combined-view learning rate comparison results at 50 epochs, with no LR scheduler and no dropout.

Based on all evaluation metrics, better prediction accuracies were observed when a learning rate scheduler was absent. A learning rate of 0.0006 at 50 epochs with no dropouts was used to examine whether or not to use an LR scheduler. The results are shown in Table 5.2

Table 5.2: Comparison between a combined-view model with a learning rate scheduler and one with no LR scheduler.

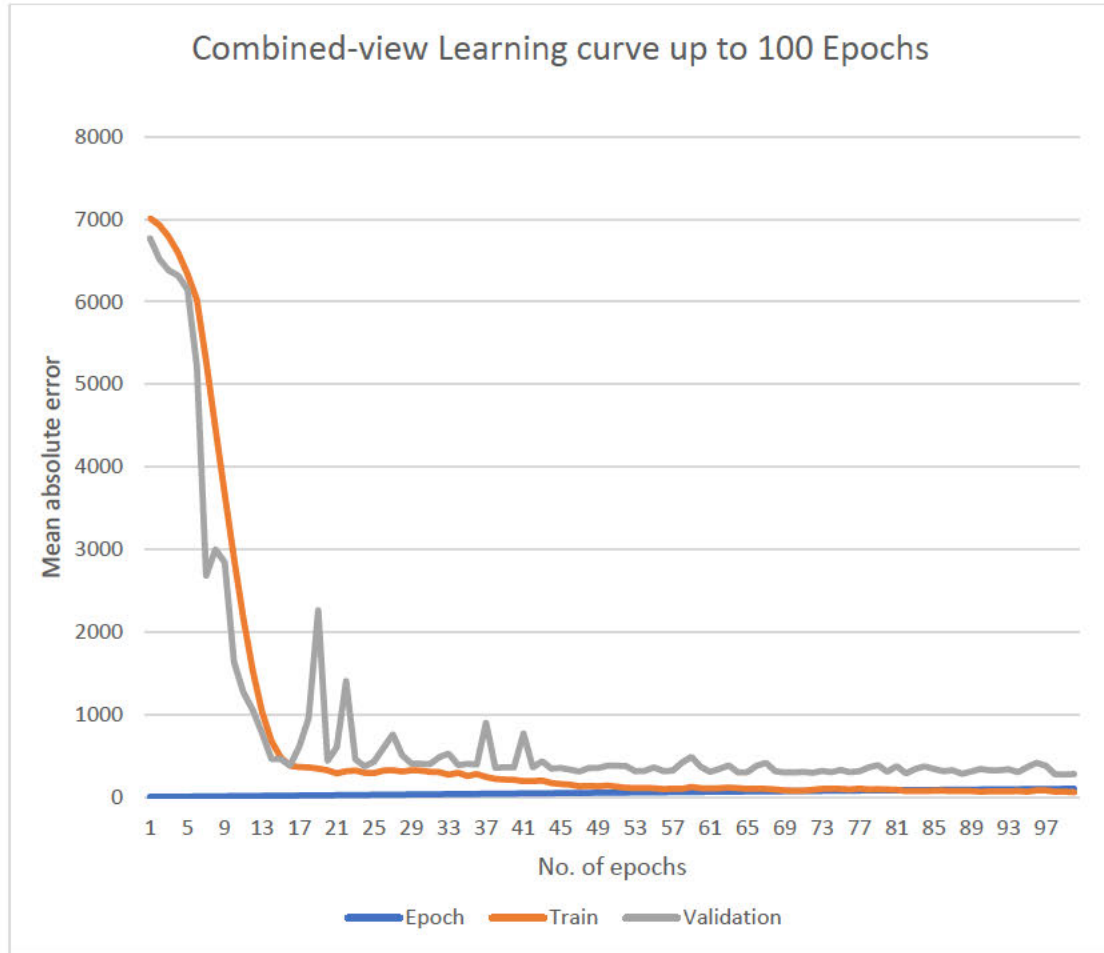
Evaluation matrix	LR Scheduler present	LR Scheduler absent
MSE	331353,874	245998,056
RMSE	575,633	495,982
MAE	418,788	312,904
R ²	0,910	0,933

LR= Learning rate, MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R²= coefficient of determination.

The Adam optimiser is the best performing optimiser for image and language processing and was used for all models in this research. The model was run at hundred epochs to establish the best number of epochs. The aim was to observe the epoch range at which the validation curve

stabilises. About fifty epochs were found to be optimum. Figure 5.4 shows the learning curves for training and validation MAEs up to 100 epochs.

Figure 5.4: Combined-view learning curves for the training and validation.



A model with a dropout rate of 0.4 was compared to one with no dropout. All evaluation matrices showed better results in models with no dropouts. The learning rate was 0.0006 and the number of epochs was 50. Table 5.3 shows the results of the comparison.

Table 5.3: effects of dropouts (dropout rate=0.4) on the combined-view model learning.

Evaluation matrix	Dropout present	Dropout absent
MSE	869026,671	245998,056
RMSE	932,216	495,982
MAE	811,623	312,904
R ²	0,763	0,933

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error and R^2 = coefficient of determination.

Concisely, a batch size of 32 and the Adam optimizer were used for all models and were determined based on literature. Based on experimental evidence, no dropouts nor LR scheduler were used for the final test. A learning rate of 0.0006 at fifty epochs was used. Table 5.4 summarises the hyperparameters used for the final combined-view model.

Table 5.4: Optimum parameter used for the final side-view analyses.

Hyperparameter	Combined-view analysis
Optimiser	Adam
Initial learning rate	0.0006
Learning rate scheduler	No
Batch size	32
The final epochs number	50
Drop out	No

5.3.6. Description of the data

Data description in terms of the sample size, amount of data used for training, validation and testing, the frequency distribution for the training and testing datasets, average milk yield, maximum MY, minimum MY and standard deviation of the data is as described in chapter 3.2.6. Both the side-view and the rear-view images were used for this analysis.

5.4. Results and Discussion

The testing data contained no augmented images and was not used in model training or validation. The model's performance dropped from the validation values of 245998.1, 496.0 kg, 312.9 kg and 0.93 for the MSE, RMSE, MAE and the R^2 value, respectively, testing values of 1963357.0, 1401.2, 1112.9, 0.17 and 0.38 for the MSE, RMSE; MAE, MAPE and the R^2 value, respectively. Table 5.5 shows the combined-view model's training, validation, and testing results at the established optimum parameters. The drop in performance from validation to testing results from the validation being done on data containing augmented images and the model testing done on unaugmented images.

Table 5.5: Training, validation, and testing results for the combined-view model with no dropouts, a learning rate of 0.0006, 50 epochs and no learning rate scheduler.

Evaluation matrix	Training	Validation	Testing
MSE	18131,985	245998,056	1963357,041
RMSE	134,655	495,982	1401,198
MAE	109,472	312,904	1112,918
MAPE			0,170
R-Square	0,995	0,933	0,375

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error, MAPE= Mean absolute percentage error and R^2 = coefficient of determination.

In all hypotheses, between the combined-view and the side-view models and between the combined-view and the rear-view models, there was no significant difference in the prediction accuracies between the models ($p>0.05$). The mean milk yield and variance for the combined view were 1168.2 kg and 904968.7, respectively. For the rear-view, the mean milk yield and variance were 1148.3 kg and 876889.9, respectively, and for the side-view, 1146.4 kg and 823121.6, respectively. The sample size was 237 for all cases. Table 5.6 summarises the two-tailed paired t-test between the combined-view model and the rear-view model, and Table 5.7 summarises the two-tailed paired t-test between the combined-view and the side-view model.

Table 5.6: Two-tailed paired t-test between the rear-view and the combined-view models.

	Rear-view	Combined view
Mean	1148,338	1168,198
Variance	876889,852	904968,660
Observations	237	237
Pearson Correlation	0,665	
Hypothesized Mean Difference	0	
df	236	
t Stat	-0,396	
P(T<=t) two-tail	0,693	
t Critical two-tail	1,970	
Results	NS	

df= degrees of freedom. NS= Not significant ($P>0.05$)

Table 5.7: Two-tailed paired t-test between the side-view and combined-view models.

	Side view	Combined view
Mean	1146,388	1168,198
Variance	823121,552	904968,660
Observations	237	237
Pearson Correlation	0,720	
Hypothesized Mean Difference	0	
df	236	
t Stat	-0,482	
P(T<=t) two-tail	0,630	
t Critical two-tail	1,970	
Results	NS	

df= degrees of freedom. NS= Not significant ($P>0.05$)

Even though using both image views to make MY predictions was expected to improve the prediction performance due to more conformation traits being considered, this did not bring significant improvement in model accuracy. The insignificant performance improvement after considering all traits is likely due to correlations between the side-view and the rear-view traits. In Chapter four, the low prediction accuracy was imputed to either the absence of traits that adequately explain milk yield variability or the method of prediction used (computer vision deep learning) and the training data size. The findings in this chapter nullify the argument provided in Chapter four that the low prediction accuracy when only one image view was considered is because of inadequate traits to explain variability in milk yield. The only commendable explanation for the poor performance is that, currently, computer vision deep learning can not be used to predict milk adequately.

Even though this analysis used more traits than Ozkaya's study, it still had a lower prediction accuracy than Ozkaya's method (Ozkaya 2015). Ozkaya (2015) used udder images to predict milk yield, and he found an R^2 value of 0.66 when all the udder traits were considered. In this study, unlike in Ozkaya's, more traits were used; factors such as the milking time, cow's standing posture, camera pixel resolution, illumination conditions, camera location and settings and distance from the cows were not standardized and deep learning was used to make predictions. Part of this study's desire was to create a model invariant to changes in factors such as the udder fullness with milk, the cow's standing posture, the distance of the cow from the camera, camera resolution and illumination of the environment. Hence these factors were not standardized in this analysis. However, this caused a reduction in prediction accuracy. Another reason for poor model performance is the full automation in this research's analysis.

Ozkaya's analysis process was more supervised in that the features to extract from the image and the trait weights were known.

The model proposed in this study had a lower accuracy than other models that used linear-type traits in MY prediction, such as the ones used by Yakubu (2011) with an R^2 of 0.69 and Gocheva-ilieva & Yordanova (2022) with an R^2 ranging from 0.93-0.95 when all traits evaluated were considered. It is important to note that Gocheva-ilieva & Yordanova (2022) also considered other factors that are not conformation traits, which were the farm, lameness and locomotion (Gocheva-ilieva & Yordanova, 2022), which, as well, improved the performance of his regression model.

The applicability of this method of milk yield prediction is not limited to Holstein cows only. It could extend to other dairy breeds such as Jersey, Ayrshire, Guernsey, Brown Swiss, and Shorthorn since linear conformation traits are visible in all these breeds.

5.4.1. Effects of splitting the test data after augmentation

A test was done to establish the effect of splitting the test data after augmentation for the combined-view model. The other parameters used in model training and validation were unchanged from the optimum parameters established in chapter 5.2.5.1. The model had an RMSE, MAE, MAPE and an R^2 of 821.3 kg, 547.1 kg, 0,09 and 0,82, respectively. based on the values shown above, the model gave better accuracies than the one where the data was split before augmentation because the model's test cases were not entirely new. This is a common data-splitting error in deep learning, as memorization in this regard does not worsen the prediction accuracy but makes it better (Shorten & Khoshgoftaar, 2019). Table 5.8 is a summary of the results of the combined-view analysis where the test data was split after image augmentation.

Table 5.8: Training, validation and testing results for the combined-view model with the test data split after augmentation, no dropouts, no learning rate scheduler, initial learning rate=0.0006 and 50 epochs.

Evaluation matrix	Training	Validation	Testing
MSE	109784,609	658653,340	674612,075
RMSE	331,338	811,575	821,348
MAE	261,588	540,199	547,121
MAPE	-	-	0,086

R-Square	0,971	0,822	0,824
----------	-------	-------	-------

MSE= Mean square error, RMSE= Root mean square error, MAE= Mean absolute error, MAPE= Mean absolute percentage error and R^2 = coefficient of determination.

Figure 5.5 is a scatter plot from the model where the test data was only split after augmentation, and Figure 5.6 is for a model where the test set was split before augmentation. The x-axis shows the predicted milk yield, and the y-axis shows actual milk yield values. The actual milk yield of a cow was also indicated by the colours on the dots, from light blue for the lowest milk yield to purple for the highest milk yield cows.

The dots in Figure 5.5 were more linear for cows below ten thousand kilograms than in Figure 5.6 where the dots are more scattered throughout. This means that for cows producing less than 10000 kgs/ lactation, predictions shown in Figure 5.5 were more accurate than those in Figure 5.6.

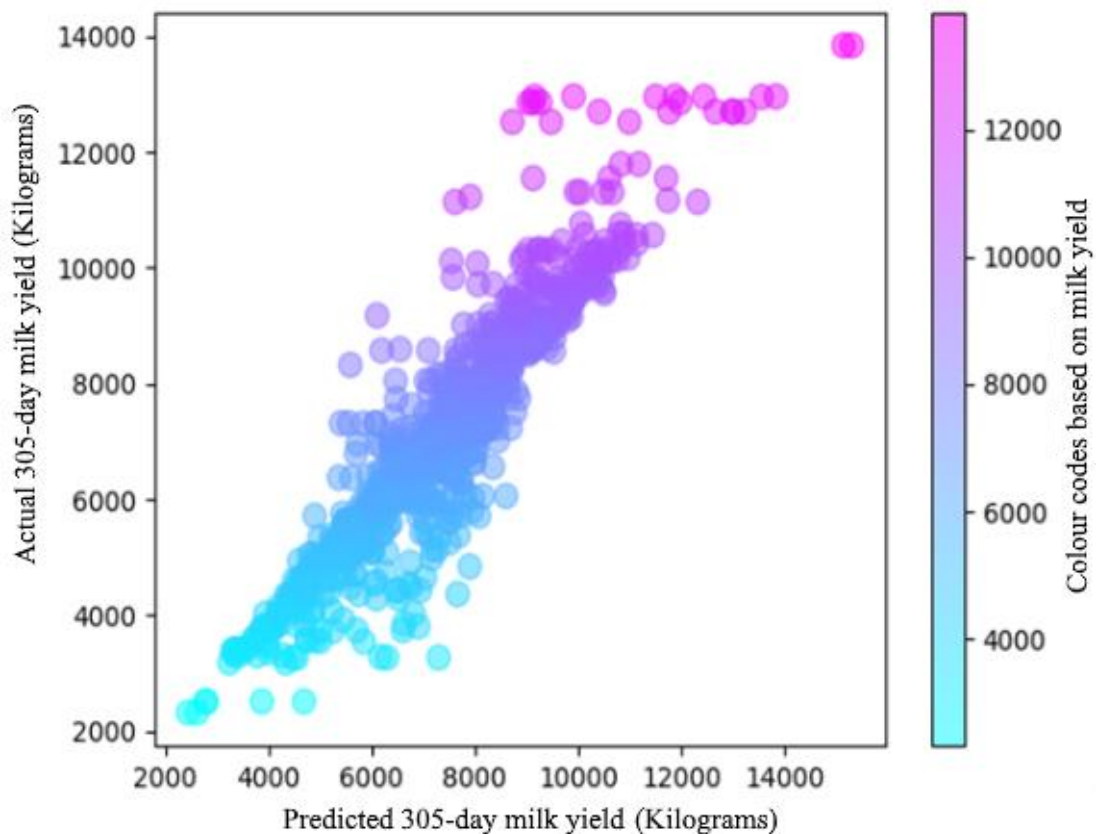


Figure 5.5: Scatter plot of the predicted vs the actual milk yield values from the combined-view model where the test data is split after augmentation.

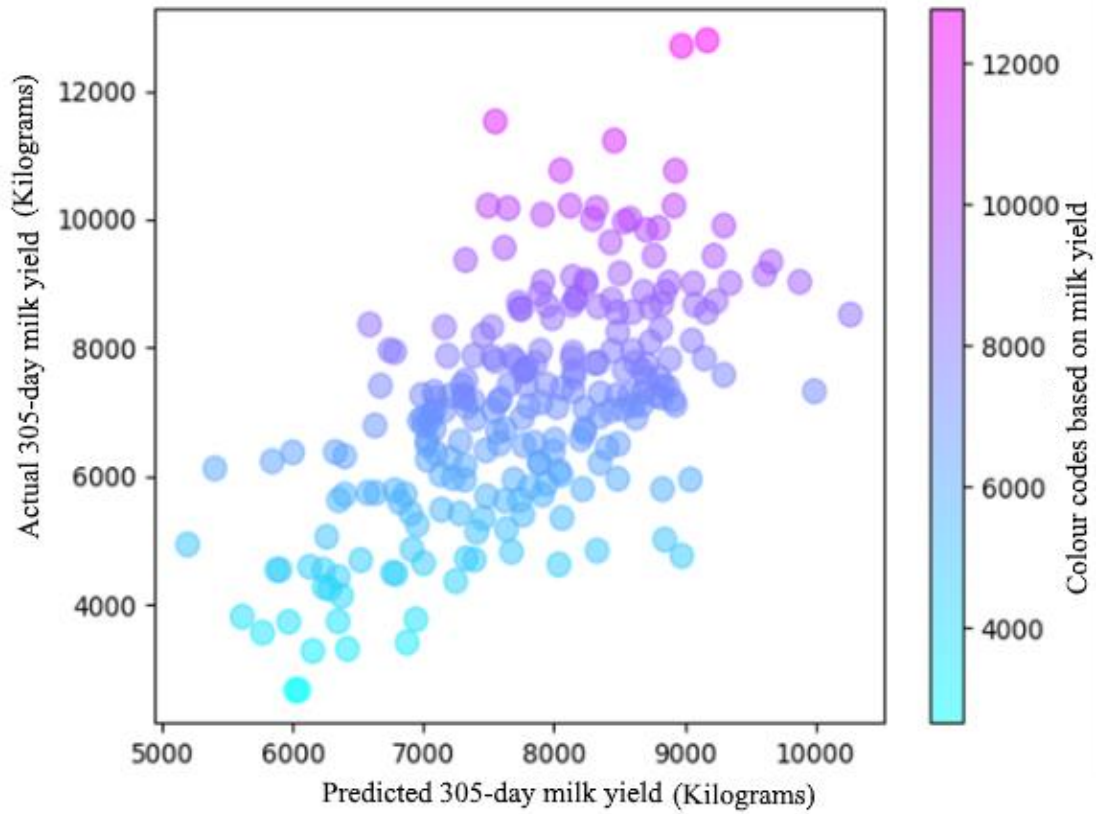


Figure 5.6: Scatter plot of the predicted vs the actual milk yield values from the combined-view model where the test data is split before augmentation.

5.5. Conclusions

Milk yield predictions can be made using cows' image data through deep learning. The predictability of 305-day milk yield using cows' combined-view images is, however weak. Also, no additional benefits were seen from using the combined-view model over models with only one image view. This suggests that either the rear or the side view has similar predictive power. Producers should use either image.

5.6. References

- Ding, X., Zhang, Z., Li, X., Wang, S., Wu, X., Sun, D., Yu, Y., Liu, J., Wang, Y., Zhang, Y., Zhang, S., Zhang, Y., & Zhang, Q. (2013). Accuracy of genomic prediction for milk production traits in the Chinese Holstein population using a reference population consisting of cows. *Journal of Dairy Science*, 96(8), 5315–5323. <https://doi.org/10.3168/jds.2012-6194>
- Gocheva-ilieva, S., & Yordanova, A. (2022). Predicting the 305-Day Milk Yield of Holstein-Friesian Cows Depending on the Conformation Traits and Farm Using Simplified Selective Ensembles. *Mathematics*. <https://doi.org/10.3390/math10081254> Academic
- Majurski, M. (2019) ‘Small Data Deep Learning : AI Applied to Domain Datasets’, Nist. Retrieved 21 August 2021 from the National Institute of Standards and technology website: https://www.nist.gov/system/files/documents/2019/08/27/workshopslides-small_data_convnets.pdf
- Ozkaya, S. (2015). Prediction possibility of milk yield from udder measurements using digital image analysis on Holstein cows. *Indian Journal of Animal Research*, 49(3), 388–391. <https://doi.org/10.5958/0976-0555.2015.00050.3>
- Shorten, C. and Khoshgoftaar, T. M. (2019) ‘A survey on Image Data Augmentation for Deep Learning’, *Journal of Big Data*, 6(1). doi: 10.1186/s40537-019-0197-0.
- Yakubu, A. (2011). Path analysis of conformation traits and milk yield of Bunaji cows in smallholder ’ s herds in Nigeria. *Agricultura Tropica et Subtropica*, 44(3), 152–157.
- Zhang, M., Luo, H., Xu, L., Shi, Y., Zhou, J., Wang, D., Zhang, X., Huang, X., & Wang, Y. (2022). Genomic Selection for Milk Production Traits in Xinjiang Brown Cattle. *Animals*, 12(2), 1–13. <https://doi.org/10.3390/ani12020136>
- Adamczak, L., Chmiel, M., Florowski, T., Pietrzak, D., Witkowski, M., & Barczak, T. (2018). The use of 3D scanning to determine the weight of the chicken breast. *Computers and Electronics in Agriculture*, 155, 394–399. <https://doi.org/10.1016/J.COMPAG.2018.10.039>
- Alphonsus, C., Akpa, G. N., Oni, O. O., Rekwot, P. I., Barje, P. P., & Yashim, S. M. (2010). Relationship of linear conformation traits with bodyweight, body condition score and milk yield in Friesian x Bunaji cows. *Journal of Applied Animal Research*, 38(1), 97–100. <https://doi.org/10.1080/09712119.2010.9707164>
- Alsahaf, A., Azzopardi, G., Ducro, B., Hanenberg, E., Veerkamp, R. F., & Petkov, N. (2019). Estimation of Muscle Scores of Live Pigs Using a Kinect Camera. *IEEE*

- Access*, 7, 52238–52245. <https://doi.org/10.1109/ACCESS.2019.2910986>
- Andonov, S., Lourenco, D. A. L., Fragomeni, B. O., Masuda, Y., Pocrnic, I., Tsuruta, S., & Misztal, I. (2017). Accuracy of breeding values in small genotyped populations using different sources of external information—A simulation study. *Journal of Dairy Science*, 100(1), 395–401. <https://doi.org/10.3168/jds.2016-11335>
- Andrew, S., Georg, H., Marc'Aurelio, R., & Ke, Y. (2013). AN EMPIRICAL STUDY OF LEARNING RATES IN DEEP NEURAL NETWORKS FOR SPEECH RECOGNITION Andrew Senior , Georg Heigold , Marc ' Aurelio Ranzato , Ke Yang New York. *New York*, 6724–6728.
- Arai, K., Kapoor, S., & Conference, C. V. (2014). Advances in Computer Vision. In *Advances in Computer Vision* (Vol. 1, Issue Cvc). <https://doi.org/10.4324/9781315802145>
- Atkins, G., Shannon, J., & Muir, B. (2008). Using Conformational Anatomy to Identify Functionality & Economics of Dairy Cows. *WCDS Advances in Dairy Technology*, 20, 279–295.
- Barnard, S., Calderara, S., Pistocchi, S., Cucchiara, R., Podaliri-Vulpiani, M., Messori, S., & Ferri, N. (2016). Quick, accurate, smart: 3D computer vision technology helps assessing confined animals' behaviour. *PLoS ONE*, 11(7). <https://doi.org/10.1371/JOURNAL.PONE.0158748>
- 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>
- Bewley, J. M., Peacock, A. M., Lewis, O., Boyce, R. E., Roberts, D. J., Coffey, M. P., Kenyon, S. J., & Schutz, M. M. (2008). Potential for Estimation of Body Condition Scores in Dairy Cattle from Digital Images. *Journal of Dairy Science*, 91(9), 3439–3453. <https://doi.org/10.3168/JDS.2007-0836>
- Bohlouli, M., Alijani, S., Varposhti, M. R., Researchers, Y., Club, E., Branch, T., Delta, A., & Company, G. (2015). *Genetic relationships among linear type traits and milk production traits of Holstein dairy cattle*. 15(4), 903–917. <https://doi.org/10.1515/aoas-2015-0053>
- Byrne, D. T., Berry, D. P., Esmonde, H., MCGovern, F., Creighton, P., & Mchugh, N. (2018). Infrared thermography as a tool to detect hoof lesions in sheep. *Translational Animal*

- Science*, 3(1), 577–588. <https://doi.org/10.1093/tas/txy132>
- Campos, R. V., Cobuci, J. A., Kern, E. L., Costa, C. N., & McManus, C. M. (2015). Genetic parameters for linear type traits and milk, fat, and protein production in Holstein cows in Brazil. *Asian-Australasian Journal of Animal Sciences*, 28(4), 476–484. <https://doi.org/10.5713/ajas.14.0288>
- Chollet, F. (2017). *Xception: Deep Learning with Depthwise Separable Convolutions*. Google, Inc.
- Cominotte, A., Fernandes, A. F. A., Dorea, J. R. R., Rosa, G. J. M., Ladeira, M. M., van Cleef, E. H. C. B., Pereira, G. L., Baldassini, W. A., & Machado Neto, O. R. (2020). Automated computer vision system to predict body weight and average daily gain in beef cattle during growing and finishing phases. *Livestock Science*, 232, 103904. <https://doi.org/10.1016/J.LIVSCI.2019.103904>
- Cross, H., Gilliland, D., Durland, P., & Seideman, S. (1983). Beef carcass evaluation by use of a video image analysis system. *Journal of Animal Sci*, 57(4), 908–917. <https://doi.org/10.2527/jas1983.574908x>
- Daetwyler, H. D., Villanueva, B., Bijma, P., & Woolliams, J. A. (2008). Inbreeding in genome-wide selection. *Journal of Animal Breeding and Genetics*, 124(6), 369–376. <https://doi.org/10.1111/j.1439-0388.2007.00693.x>
- Ding, X., Zhang, Z., Li, X., Wang, S., Wu, X., Sun, D., Yu, Y., Liu, J., Wang, Y., Zhang, Y., Zhang, S., Zhang, Y., & Zhang, Q. (2013). Accuracy of genomic prediction for milk production traits in the Chinese Holstein population using a reference population consisting of cows. *Journal of Dairy Science*, 96(8), 5315–5323. <https://doi.org/10.3168/jds.2012-6194>
- Doeschl, A. B., Green, D. M., Whittemore, C. T., Schofield, C. P., Fisher, A. V., & Knap, P. W. (2004). The relationship between the body shape of living pigs and their carcass morphology and composition. *Animal Science*, 79(1), 73–83. <https://doi.org/10.1017/s1357729800054540>
- Ehrlich, J. L. (2011). Quantifying shape of lactation curves , and benchmark curves for common dairy breeds and parities. *The Bovine Practitioner*, 45(1), 88–96. http://aabp.org/members/publications/2011/prac_feb/16-Ehrlich.pdf
- FAO. (2021). Dairy market review, Food and Agriculture Organisation of the United Nations. *Food and Agriculture Organization of the United Nations, April*, 1–13.

- Fernandes, Arthur F A, Dórea, J. R. R., Fitzgerald, R., Herring, W., & Rosa, G. J. M. (2019). A novel automated system to acquire biometric and morphological measurements and predict body weight of pigs via 3D computer vision 1. *J. Anim. Sci*, *97*, 496–508. <https://doi.org/10.1093/jas/sky418>
- Fernandes, Arthur F A, Dórea, J. R. R., Valente, B. D., Fitzgerald, R., Herring, W., & Rosa, G. J. M. (2020). Comparison of data analytics strategies in computer vision systems to predict pig body composition traits from 3D images. *Journal of Animal Science*, *98*(8). <https://doi.org/https://doi.org/10.1093/jas/skaa250>
- Fernandes, Arthur Francisco Araújo, Dórea, R. J. R., & Rosa, G. J. D. M. (2020). Image Analysis and Computer Vision Applications in Animal Sciences: An Overview. *Frontiers in Veterinary Science*, *7*(October), 1–18. <https://doi.org/10.3389/fvets.2020.551269>
- Forni, S., Aguilar, I., & Misztal, I. (2011). Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genetics Selection Evolution*, *43*(1), 1–7. <https://doi.org/10.1186/1297-9686-43-1>
- Getu, A., & Misganaw, G. (2015). The Role of Conformational Traits on Dairy Cattle Production and Their Longevities. *OALib*, *02*(03), 1–9. <https://doi.org/10.4236/oalib.1101342>
- Gocheva-ilieva, S., & Yordanova, A. (2022). Predicting the 305-Day Milk Yield of Holstein-Friesian Cows Depending on the Conformation Traits and Farm Using Simplified Selective Ensembles. *Mathematics*. <https://doi.org/10.3390/math10081254> Academic
- Gorgulu, O. (2012). Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. *South African Journal of Animal Science*, *42*(3). <https://doi.org/10.4314/sajas.v42i3.10>
- Grzesiak, W., Wojcik, J., & Binerowska, B. (2003). Prediction of 305-day first lactation milk yield in cows with selected regression models. *Archives Animal Breeding*, *46*(3), 215–226. <https://doi.org/10.5194/aab-46-213-2003>
- Habier, D., Fernando, R. L., & Dekkers, J. C. M. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, *177*(4), 2389–2397. <https://doi.org/10.1534/GENETICS.107.081190>
- Halachmi, I., Klopčič, M., Polak, P., Roberts, D. J., & Bewley, J. M. (2013). Automatic assessment of dairy cattle body condition score using thermal imaging. *Computers and Electronics in Agriculture*, *99*, 35–40. <https://doi.org/10.1016/J.COMPAG.2013.08.012>

- Harris, R. (2015). *Phenotypic correlations between linear type conformation traits, production and fertility in a once-a-day milked dairy cattle herd: a thesis presented in partial fulfillment*. <http://mro.massey.ac.nz/handle/10179/7430>
- Hassaballah, M., & Awad, A. I. (2020). Deep Learning in Computer Vision. In *Deep Learning in Computer Vision* (Issue March). <https://doi.org/10.1201/9781351003827>
- Hordri, N. F., Samar, A., Yuhaniz, S. S., & Shamsuddin, S. M. (2017). A systematic literature review on features of deep learning in big data analytics. *International Journal of Advances in Soft Computing and Its Applications*, 9(1), 32–49.
- Hovinen, M., Siivonen, J., Taponen, ‡ S, Hänninen, L., Pastell, † M, Aisla, A.-M., & Pyörälä, S. (2008). Detection of Clinical Mastitis with the Help of a Thermal Camera. *Journal of Dairy Science*, 91, 4592–4598. <https://doi.org/10.3168/jds.2008-1218>
- ICAR. (2015). *1. Conformation Recording of Dairy Cattle. June*, 1–42. <https://www.icar.org/wp-content/uploads/2015/08/Conformation-Recording-CR-WG.pdf>
- International Committee for Animal Recording. (2010). *Conformation recording of dairy cattle. May*, 1–20.
- Janković, D., Trivunović, S., & Samolovac, L. (2020). *Characteristics of linear traits of udder and angularity in Holstein-friesian cows and their correlation with milk yield traits*. 36(4), 407–416.
- Jensen, D. B., van der Voort, M., & Hogeveen, H. (2018). Dynamic forecasting of individual cow milk yield in automatic milking systems. *Journal of Dairy Science*, 101(11), 10428–10439. <https://doi.org/10.3168/JDS.2017-14134>
- Jeretina, J., Babnik, D., & Škorjanc, D. (2016). Prediction of Standard Lactation Curves for Primiparous Holstein Cows by Using Corrected Regression Models Prediction of standard lactation curves for primiparous Holstein cows by using corrected regression models. *Italian Journal of Animal Science*. <https://doi.org/10.4081/ijas.2015.3776>
- Jingar, S., Mehla, R. K., Singh, M., & Roy, A. K. (2014). Lactation Curve Pattern and Prediction of Milk Production Performance in Crossbred Cows. *Journal of Veterinary Medicine*, 2014, 1–6. <https://doi.org/10.1155/2014/814768>
- Kandel, I., & Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, 6(4), 312–315. <https://doi.org/10.1016/j.ict.2020.04.010>

- Kashiha, M., Bahr, C., Haredasht, S. A., Ott, S., Moons, C. P. H., Niewold, T. A., Ödberg, F. O., & Berckmans, D. (2013). The automatic monitoring of pigs water use by cameras. *Computers and Electronics in Agriculture*, *90*, 164–169. <https://doi.org/10.1016/J.COMPAG.2012.09.015>
- Kern, E. L., Cobuci, J. A., Costa, C. N., McManus, C. M., Campos, G. S., & Almeida, T. P. (2015). Genetic association between longevity and linear type traits of holstein cows. *Scientia Agricola*, *72*(3), 203–209. <https://doi.org/10.1590/0103-9016-2014-0007>
- Khan, M. S. M. A., & Khan, M. S. M. A. (2016). Genetic and phenotypic correlations between linear type traits and milk yield in Sahiwal cows. *Pakistan Journal of Agricultural Sciences*, *53*(2), 483–489. <https://doi.org/10.21162/PAKJAS/16.3369>
- Khoshnoudi-Nia, S., & Moosavi-Nasab, M. (2019). Prediction of various freshness indicators in fish fillets by one multispectral imaging system. *Scientific Reports*, *9*(1). <https://doi.org/10.1038/S41598-019-51264-Z>
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.
- Kliś, P., Piwczyński, D., Sawa, A., & Sitkowska, B. (2021). Prediction of lactational milk yield of cows based on data recorded by AMS during the periparturient period. *Animals*, *11*(2), 1–11. <https://doi.org/10.3390/ani11020383>
- Kongsro, J. (2014). Estimation of pig weight using a Microsoft Kinect prototype imaging system. *Computers and Electronics in Agriculture*, *109*, 32–35. <https://doi.org/10.1016/J.COMPAG.2014.08.008>
- Li, N., Ren, Z., Li, D., & Zeng, L. (2020). Review: Automated techniques for monitoring the behaviour and welfare of broilers and laying hens: towards the goal of precision livestock farming. *Animal*, *14*(3), 617–625. <https://doi.org/10.1017/S1751731119002155>
- Liseune, A., Salamone, M., Van den Poel, D., Van Ranst, B., & Hostens, M. (2020). Leveraging latent representations for milk yield prediction and interpolation using deep learning. *Computers and Electronics in Agriculture*, *175*(July), 105600. <https://doi.org/10.1016/j.compag.2020.105600>
- Lucas, D., Brun, A., Gispert, M., Carabús, A., Soler, J., Tibau, J., & Font-i-Furnols, M. (2017). Relationship between pig carcass characteristics measured in live pigs or carcasses with Piglog, Fat-o-Meat'er and computed tomography. *Livestock Science*,

197, 88–95. <https://doi.org/10.1016/J.LIVSCI.2017.01.010>

Luis Nunes, J., Piquerez, M., Pujadas, L., Armstrong, E., Fernández, A., & Lecumberry, F. (2015). *Beef quality parameters estimation using ultrasound and color images*.

<https://doi.org/10.1186/1471-2105-16-S4-S6>

Martins, R. F. S., do Prado Paim, T., de Abreu Cardoso, C., Stéfano Lima Dallago, B., de Melo, C. B., Louvandini, H., & McManus, C. (2013). Mastitis detection in sheep by infrared thermography. *Research in Veterinary Science*, 94(3), 722–724.

<https://doi.org/10.1016/J.RVSC.2012.10.021>

Matthews, S. G., Miller, A. L., Plötz, T., & Kyriazakis, I. (2017). Automated tracking to measure behavioural changes in pigs for health and welfare monitoring. *Scientific Reports*, 7(1), 1–12. <https://doi.org/10.1038/s41598-017-17451-6>

Metzner, M., Sauter-Louis, C., Seemueller, A., Petzl, W., & Zerbe, H. (2015). Infrared thermography of the udder after experimentally induced *Escherichia coli* mastitis in cows. *The Veterinary Journal*, 204(3), 360–362.

<https://doi.org/10.1016/J.TVJL.2015.04.013>

Nasirahmadi, A., Edwards, S. A., & Sturm, B. (2017). Implementation of machine vision for detecting behaviour of cattle and pigs. *Livestock Science*, 202, 25–38.

<https://doi.org/10.1016/J.LIVSCI.2017.05.014>

Nguyen, Q. T., Fouchereau, R., Frénod, E., Gerard, C., & Sincholle, V. (2020). Comparison of forecast models of production of dairy cows combining animal and diet parameters. *Computers and Electronics in Agriculture*, 170.

<https://doi.org/10.1016/j.compag.2020.105258>

Njubi, D. M., Wakhungu, J. W., & Badamana, M. S. (2010). Use of test-day records to predict first lactation 305-day milk yield using artificial neural network in Kenyan Holstein-Friesian dairy cows. *Tropical Animal Health and Production*, 42(4), 639–644.

<https://doi.org/10.1007/s11250-009-9468-7>

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G. V., Krpalkova, L., Riordan, D., & Walsh, J. (2020). Deep Learning vs. Traditional Computer Vision. *Advances in Intelligent Systems and Computing*, 943(Cv), 128–144.

https://doi.org/10.1007/978-3-030-17795-9_10

Ogundeji, A. A., Lakew, H., Tesfahuney, W., & Lombard, W. (2021). Influence of heat stress on milk production and its financial implications in semi-arid areas of South Africa. *Heliyon*, 7(5), e06202. <https://doi.org/10.32964/tj19.5>

- Ozkaya, S. (2015). Prediction possibility of milk yield from udder measurements using digital image analysis on Holstein cows. *Indian Journal of Animal Research*, 49(3), 388–391. <https://doi.org/10.5958/0976-0555.2015.00050.3>
- Paluchowski, L. A., Misimi, E., Grimsmo, L., & Randeberg, L. L. (2016). Towards automated sorting of Atlantic cod (*Gadus morhua*) roe, milt, and liver – Spectral characterization and classification using visible and near-infrared hyperspectral imaging. *Food Control*, 62, 337–345. <https://doi.org/10.1016/J.FOODCONT.2015.11.004>
- Patel, K., Bancroft, N., Drucker, S. M., Fogarty, J., Ko, A. J., & Landay, J. A. (2010). Gestalt: Integrated support for implementation and analysis in machine learning. *UIST 2010 - 23rd ACM Symposium on User Interface Software and Technology*, 37–46. <https://doi.org/10.1145/1866029.1866038>
- Patil, A., & Rane, M. (2021). Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition. *Smart Innovation, Systems and Technologies*, 195, 21–30. https://doi.org/10.1007/978-981-15-7078-0_3
- Pezzuolo, A., Guarino, M., Sartori, L., González, L. A., & Marinello, F. (2018). On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera. *Computers and Electronics in Agriculture*, 148, 29–36. <https://doi.org/10.1016/J.COMPAG.2018.03.003>
- Phahlane, H., Mthembeka, Z. A., Lekganyane, M. S., Mofolo, K. M., Gininda, P., Ramonyai, D., & Lubbe, P. (2021). Quaterly economic overview. *The Department of Agriculture, Forestry and Fisheries, Volume 19*, 5–24.
- Psota, E. T., Mittek, M., Pérez, L. C., Schmidt, T., & Mote, B. (2019). Multi-pig part detection and association with a fully-convolutional network. *Sensors (Switzerland)*, 19(4), 1–24. <https://doi.org/10.3390/s19040852>
- Saberioon, M., Gholizadeh, A., Cisar, P., Pautsina, A., & Urban, J. (2016). *Application of machine vision systems in aquaculture with emphasis on fish: state-of-the-art and key issues*. <https://doi.org/10.1111/raq.12143>
- Sassi, N. Ben, Averós, X., Estevez, I., Nicol, C., & Rodenburg, T. B. (2016). *animals Technology and Poultry Welfare*. <https://doi.org/10.3390/ani6100062>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0197-0>

- Smith, G. (2021). South Africa's dairy sector strives to shake off its past. In *The Daily Churn*. Darigold. <https://www.darigold.com/south-africas-dairy-sector-strives-to-shake-off-its-past/>
- South African milk processors' organization. (2020). Summary of Key Market Signals for the Dairy Industry , November 2020 Edition. *Annual Report, November 2020*.
- South African Milk Processors' Organization. (2021). Summary of Key Market Signals for the Dairy Industry , August 2021 Edition. *Annual Report, August 2021*.
- 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>
- Stamschror, J., & Card, S. (2000). Judging Dairy Cattle. *System*, July, 1–8.
- Tapki, I. (2013). Genetic and Phenotypic Correlations between Linear Type Traits and Milk Production Yields of Turkish Holstein Dairy Cows By. *Green Journal of Agricultural Sciences*, 3(11), 755–761. <https://doi.org/10.15580/GJAS.2013.11.072913763>
- Vallejo, R. L., Leeds, T. D., Gao, G., Parsons, J. E., Martin, K. E., Evenhuis, J. P., Fragomeni, B. O., Wiens, G. D., & Palti, Y. (2017). Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genetics Selection Evolution*, 49(1), 1–13. <https://doi.org/10.1186/s12711-017-0293-6>
- Waal, H. De, & Blom, C. (2020). Annual Milk Cattle Bulletin. *National Milk Recording and Improvement Scheme*.
- Wall, E., White, I. M. S., Coffey, M. P., & Brotherstone, S. (2005). The relationship between fertility, rump angle, and selected type information in Holstein-Friesian cows. *Journal of Dairy Science*, 88(4), 1521–1528. [https://doi.org/10.3168/jds.S0022-0302\(05\)72821-6](https://doi.org/10.3168/jds.S0022-0302(05)72821-6)
- Wang, X., Zhao, M., Ju, R., Song, Q., Hua, D., Wang, C., & Chen, T. (2013). Visualizing quantitatively the freshness of intact fresh pork using acousto-optical tunable filter-based visible/near-infrared spectral imagery. *Computers and Electronics in Agriculture*, 99, 41–53. <https://doi.org/10.1016/j.compag.2013.08.025>
- Wassenberg, R., Allen, D., & Kemp, K. (1986). Video image analysis prediction of total kilograms and percent primal lean and fat yield of beef carcasses. *Journal of Animal Sci*, 62(6), 1609–1616. <https://doi.org/10.2527/jas1986.6261609x>

- Yakubu, A. (2011). Path analysis of conformation traits and milk yield of Bunaji cows in smallholder ' s herds in Nigeria. *Agricultura Tropica et Subtropica*, 44(3), 152–157.
- Zapotoczny, P., Szczypiński, P. M., & Daszkiewicz, T. (2016). Evaluation of the quality of cold meats by computer-assisted image analysis. *LWT - Food Science and Technology*, 67, 37–49. <https://doi.org/10.1016/J.LWT.2015.11.042>
- Zhang, F., Shine, P., Upton, J., Shaloo, L., & Murphy, M. D. (2018). A review of milk production forecasting models: Past and future methods. *Dairy Farming: Operations Management, Animal Welfare and Milk Production, January 2019*, 13–61.
- Zhang, M., Luo, H., Xu, L., Shi, Y., Zhou, J., Wang, D., Zhang, X., Huang, X., & Wang, Y. (2022). Genomic Selection for Milk Production Traits in Xinjiang Brown Cattle. *Animals*, 12(2), 1–13. <https://doi.org/10.3390/ani12020136>
- Zindove, T. J., Chimonyo, M., & Nephawe, K. A. (2014). Relationship between linear type and fertility traits in Nguni cows. *Animal*, 9(6), 944–951. <https://doi.org/10.1017/S1751731114003231>
- Zink, V., Zavadilová, L., Lassen, J., Štípková, M., Vacek, M., & Štolc, L. (2014). Analyses of genetic relationships between linear type traits, fat-to-protein ratio, milk production traits, and somatic cell count in first-parity Czech Holstein cows. *Czech Journal of Animal Science*, 59(12), 539–547. <https://doi.org/10.17221/7793-cjas>

CHAPTER SIX: General discussion, Conclusions and Recommendations

6.1. General discussion

Three hypotheses were tested, milk yield prediction is the same when using either the side-view or the rear-view images of cows, milk yield prediction is the same when using either the combined-view or only the rear-view images of cows, and milk yield prediction is the same when using either the combined-view or only the side-view images of cows. The hypotheses were formulated based on previous findings where milk yield prediction was made using only the rear-view traits (Ozkaya, 2015) and where milk yield prediction was made using conformation traits from all over the cow's body (Yakubu, 2011; Gocheva-ilieva & Yordanova, 2022). In all the hypothesis tests, no significant difference was observed between the prediction accuracies of the model using the side-view, the rear-view and the combined-view images.

Through deep learning, chapter 3 determined the predictability of 305-day milk yield using cows' rear-view images. The rear-view image contained these essential traits: rump width, rear udder width, udder depth, rear udder height, teat placement and length. The Xception architecture (a Convolutional Neural Network model) was used as the deep learning model. The predictability of milk yield based on the rear-view traits was found to be poor ($R^2=0.30$ and $MAPE=0.17$). The poor predictability of 305- day milk yield using cows' rear-view images was attributed mainly to the method of analysis, i.e. computer vision deep learning and not the linear conformation traits. This is because other research has proved it possible to predict milk yield from rear-view images of cows (Ozkaya, 2015).

Chapter 4 assessed the predictability of the 305-day milk yield based on the side-view images through deep learning. The captured traits from the side-view images were angularity, fore udder attachment, side-view teat placement and length, rump angle and udder depth. Through the Xception network, the predictability of milk yield based on the side-view images was poor ($R^2= 0.32$ and $MAPE=0.18$). This model's performance was not different from the model using the rear-view images, suggesting that traits seen from the side-view are wholistically as good at predicting MY as those seen from the posterior of the cow. Again the poor predictability of 305- day milk yield using cows' side-view images was ascribed to the prediction method used and the full automation thereof.

The final analysis was for the predictability of milk yield using the combined-view through computer vision deep learning. This was presented in chapter 5. This model incorporated all traits seen from the side and those seen from the posterior of the cow. The prediction accuracy

of the model was generally poor ($R^2=0.38$ and $MAPE= 0.17$). The poor performance was also imputed mainly to the model used as there were other methods of prediction based on linear-type traits that gave better predictions (Yakubu, 2011; Gocheva-ilieva & Yordanova, 2022). Irrespective of the increase in the conformation traits used, the model accuracy was not different from the other two models where only the side-view or the rear-view was considered in making MY predictions. This was likely because of the correlations between the side-view and the rear-view linear-type traits.

For all analyses, predictions were weaker for very high-producing cows, i.e. $MY>10000$ Kgs/lactation. This could be a result of a limited number of cows within this range for proper model learning or because there are little linear trait differences between high-producing and very high-producing cows. Other factors besides linear conformation traits may be impelling a cow to move from high milk production to very high milk production. These could be factors such as management, feed quality and availability and other genetic factors such as feed conversion efficiency.

These studies can be implemented on other dairy breeds such as Jersey, Guernsey, Ayrshire, Brown Swiss, and Shorthorn. Deep learning models specific to these breeds can be made to predict milk yield.

6.2. Conclusions

Predicting the 305-day milk yield of Holstein cows using either the side or the rearview has the same level of accuracy. No additional benefits are derived from using a combination of the rear and the side view images. The predictions are, however, weak for all image views, the rear-view, the side-view and a combination of the rear-view and the side-view. The 305-day milk yield prediction was poorer for cows producing milk above ten thousand kilograms.

6.3. Recommendations

It is recommended that aspects of the images such as cow distance from the camera, cow postures, udder fullness and environmental illumination be standardized to maximise prediction performance through image data. This, however, will narrow down the types of images to which milk yield prediction can be made.

Feather research is required on the effects of feature extraction before machine learning on milk yield prediction using the same data conditions. Feature extraction will also be salient for establishing the actual rear-view or side-view traits that the model uses to make predictions and the weights each trait carries. It may also be necessary to try out other CNN architectures

such as ResNet, EfficientNet, AlexNet or VGGNet on the same dataset to establish whether any of these networks create a better prediction model for MY.

6.4. References

- Gocheva-ilieva, S., & Yordanova, A. (2022). Predicting the 305-Day Milk Yield of Holstein-Friesian Cows Depending on the Conformation Traits and Farm Using Simplified Selective Ensembles. *Mathematics*. <https://doi.org/10.3390/math10081254> Academic
- Ozkaya, S. (2015). Prediction possibility of milk yield from udder measurements using digital image analysis on Holstein cows. *Indian Journal of Animal Research*, 49(3), 388–391. <https://doi.org/10.5958/0976-0555.2015.00050.3>
- Yakubu, A. (2011). Path analysis of conformation traits and milk yield of Bunaji cows in smallholder's herds in Nigeria. *Agricultura Tropica et Subtropica*, 44(3), 152–157.

Appendix 1: Data collection approval letter from The Glen farm



CC REG. NO. 2007/019949/23 - Member: Johnny van Rensburg

Die Glen, P.O. Box 239, Humansdorp 6300

Tel: 042 291 0477, Cell: 082 498 0566, Fax: 086 582 5786, Email: groenwei@igen.co.za

25 October 2021

Mr Lawrence Jembere
MScAG in Animal Science
2 Amersham road
Westville
DURBAN
3629

Dear Mr L Jembere

Re: Permission to collect data for the research: Prediction of milk yield using cows' visual images through Deep Learning

We are pleased to inform you that permission regarding your request to collect data for our research at The Glen Dairy farm was granted. We properly read your proposed data collection protocols, and there appears to be no animal misconduct or mishandling. In case of any unanticipated event, while taking the pictures, you will be required to report them to any of the management team immediately. Data on milk yield, parity, age and BCS for the cows will be emailed to you as per your request.

Best wishes on your research

Yours sincerely



JJ van Rensburg

**SANDWATER BOERDERY
TRUST**

Posbus 342
Humansdorp, 6300
Tel: (042) 2951760
VAT: 457 016 9427

MScAG in Animal Science
2 Amersham road
Westville
Durban
3629

20 October 2021

Dear Mr L. Jembere

Re: Permission to collect data for the research: Prediction of milk yield using cows' visual images through Deep Learning

We are pleased to inform you that permission regarding your request to collect data for your research at Sandwater Dairy farm was granted. We properly read your proposed data collection protocols, and there appears to be no animal misconduct or mishandling of animals. In case of any unanticipated event while taking the pictures, you will be required to report them to any of the management team immediately. Data on milk yield, parity, age and BCS for the cows will be emailed to you as per your request.

Sandwater is a registered commercial dairy farm and is currently not under any disease restrictions. We will communicate with you in case of any disease outbreak during the data collection period.

Best wishes on your research

Yours sincerely


Izak van der Merwe

ANDREW MASTERSON FAMILY TRUST

Milagro Farm, Oyster Bay
PO Box 117, Humansdorp, 6300
Tel 042 291 0203 Fax: 086 589 4338
VAT no: 4380191009

11 April 2022

To whom it may concern

Re: Research and data collection by Lawrence Jembere

I hereby confirm that Lawrence Jembere can collect data from my Holstein herd on farm and I hereby give authorisation for him to use this data in his research project.

Name: Lawrence Jembere

University: University of KwaZulu-Natal

Student number: 221089192

Program: MScAG in Animal Science

Email: 221089192@stu.ukzn.ac.za

I trust that you will find this in order.

Yours sincerely



Andrew Masterson
Cell: 082 3211 462
Email: andrew@milagro.co.za



Registration no: IT 1420/2000/2

BLOEKOMLAAN AGRI (PTY)LTD

VAT NO. 4890268248
REG NO IT 2014/076897/07
PO BOX 423, HUMANSDORP, 6300
TEL NO. 087 700 8365
FAX NO. 086 211 6852

To Whom It May Concern

Subject: Approval Letter for Data Collection (MScAG in Animal Science)

I Chris Du Plooy, Manager of Bloekomlaan Dairy,

granted student Lawrence Jembere, Student nr: 221089192, permission to collect Data for his research topic; Prediction of milk yield using cow's visual images through deep learning.

It is our pleasure to accommodate him on the farm and hope he will gain knowledge (mention the topic) from your data collected. I can assure you that information will be given from our side when needed. We will be pleased if our data will help you in your research work.

If you have any queries or requests, feel free to contact me in this regard.

Best of luck for your studies and future.

Kind Regards



Chris Du Plooy

BLOEKOMLAAN AGRI (PTY)LTD
VAT NO. 4890268248
REG NO IT 2014/076897/07
PO BOX 423, HUMANSDORP, 6300
TEL NO. 087 700 8365
EMAIL jf@bomhunt.co.za