International Conference of Agricultural Engineering
AgEng 2014 Zurich 6–10 July

Ref: C0685

Improving a computer vision lameness detection system by adding behaviour and performance measures

Tom Van Hertem, Machteld Steensels, Stefano Viazzi, Eduardo C. B. Romanini, Claudia Bahr and Daniel Berckmans, Division M3-BIORES: Measure, Model & Manage Bioresponses, Katholieke Universiteit Leuven, Kasteelpark Arenberg 30 – bus 2456, BE-3001 Heverlee, Belgium

Andres Schlageter Tello and Kees Lokhorst, Wageningen UR Livestock Research, P.O. Box 65, NL-8200AB Lelystad, the Netherlands

Ephraim Maltz and Ilan Halachmi, Institute of Agricultural Engineering – Agricultural Research Organization (ARO) – the Volcani Center, P.O. Box 6, IL-50250 Bet Dagan, Israel

Abstract

Monitoring cow behavior and performance is a key factor in dairy health and welfare management. Due to increased farm size and limited time per animal, visual observation of cow health and welfare is evolving to automated monitoring systems. Computer vision is a promising technique for cow monitoring because it is relatively low cost and simple to install. The spread in use of behaviour and performance sensors make continuous monitoring of the individual in the herd possible The hypothesis was that adding available data such as activity and milk yield to a computer vision based lameness detection model would increase the classification accuracy. The aim of this study was to prove that the addition data from behaviour and performance sensor sources improved the lameness classification accuracy of the lameness-specific feature variables extracted from recorded videos.
Data were gathered on a commercial dairy farm in Belgium. Between 1 September 2013 and 30 April 2014, 40 different locomotion score sessions were performed on the farm. Cow gait was recorded automatically in the corridor when the cows returned from the rotary milking parlour to the cowshed. All recordings were made with a depth camera. The recorded videos were processed by a software algorithm that extracted image feature variables (theta1, theta2, theta3, Back Posture Measurement, inverse radius, walking speed). The complete image acquisition and image processing was done automatically.
The neck collar of all cows was equipped with a commercial tag that measured cow activity in one hour intervals. Cows were milked twice per day, and milk yield, milk conductivity, milk flow rate and milking order were measure at each individual milking stand. All data were standardized to daily values.
Classification performance was quantified by the area under the receiver operating characteristics curve (AUC). Univariate models were better for video-related variables (AUC = [0.5452 – 0.7199]) than for activity-related (AUC = [0.5397 – 6155]) and milking-related (AUC = [0.5359 – 0.5846]) variables. The forward stepwise binary logistic regression resulted in a multivariate model with 10 input variables (walking speed, theta1, BPM, daily activity, daytime activity, daily milk yield, milk conductivity, milking order, lactation stage and milk peak flow rate). This study shows that the multivariate models from different sensor input data sources give better classification results than univariate classification models.

Keywords: automation, locomotion, computer vision, behaviour sensing, back posture

1. Introduction

Lameness is a one of the biggest health and welfare issue in modern intensive dairy farming (Bruijnis, Beerda, Hogeveen, & Stassen, 2012). The prevalence of lameness is often underestimated, and is affected by many different factors. The most common method to obtain a herd lameness prevalence rate is visual locomotion scoring (Flower, & Weary, 2009). This procedure is however subjective, time-consuming and costly. Monitoring cow behavior and performance is a key factor in dairy health and welfare management. Due to increased farm size and limited time per animal, visual observation of cow health and welfare is evolving to automated monitoring systems. Computer vision is a promising technique for cow monitoring because it is relatively low cost and simple to install. Studies using computer vision for gait analysis focus on back arch curvature (Poursaberi, Bahr, Pluk, Van Nuffel, & Berckmans, 2010; Viazzi et al., 2014; Van Hertem et al., 2014), step overlap (Pluk et al., 2010), and hoof release angles (Pluk et al., 2012). Back curvature can be extracted from images by different feature variables, such as an inverse radius (Poursaberi, Bahr, Pluk, Van Nuffel, & Berckmans, 2010; Viazzi et al., 2014), different angles (Viazzi et al., 2014) and a Back Posture Measurement (BPM) (Van Hertem et al., 2014). All extracted feature variables show a strong relation to locomotion scores. However, implementing a computer vision systems in commercial farm conditions tends to be challenging.
Automatic milk yield measurements are very common in European dairy farming, and the use of other sensors such as activity sensors, milk conductivity sensors etc. is spreading (Rutten, Velthuis, Steeneveld, & Hogeveen, 2013). Automatic sensors make continuous monitoring of the individual in the herd possible. However, the large pool of data coming from these sensors is not extensively used. Estrus detection based on animal behaviour is well-described in literature (Jonsson, Blanke, Poulsen, Caponetti, & Hojsgaard, 2011) and found its way to the commercial market. There are also studies on automatic mastitis detection (Huybrechts, Mertens, De Baerdemaeker, De Ketelaere, & Saeys, 2014) and lameness detection (Van Hertem et al., 2013). The use of a single predictor variable was very often not strong enough as a classifier, partly due to high within- and between-cow variability (Kramer, Cavero, Stamer, & Krieter, 2009), but it was often suggested to use in a multi-pronged approach (Ito, von Keyserlingk, LeBlanc, & Weary, 2010).
To our knowledge, no attempt was done to combine computer vision data with other sensor data in order to reach a better diagnostic tool for lameness. The hypothesis was that adding more frequent available data to a lameness detection model would increase the classification accuracy. The aim of this study was to prove that the addition of behaviour and performance data improved the lameness classification accuracy of the lameness specific feature variables extracted from videos.

2. Materials and methods

2.1. Animals and housing

All data were gathered from a commercial dairy farm in Arendonk, Belgium in the period 1 September 2013 till 30 April 2014. The number of cows in the milking herd ranged between 210 and 240. All cows were from the Holstein-Friesian breed and were housed all-year-round indoors in a cubicle barn with slatted floors. The milking herd was divided in 2 production groups according to production level, and the proportional group distribution was on average (high)3:2(low). The cows were milked two times per day ([06.00h – 08.30h] and [18.00h – 20.15h]) in a 40-stand DeLaval rotary milking parlour. Prior to milking, both production groups were brought to the waiting area. An automated mechanic fence brought the cows closer to the rotary. After milking, the cows stepped away from the rotary milking platform, and entered a 20m long corridor/alley that brought them back to the cow shed. At the end of the corridor, a spray box disinfected the udder and teats after milking, and a smart selection gate regulated cow traffic.

2.2. Video data acquisition

Cow gait recordings were made with a 3D image camera (Kinect™, Microsoft corp., Redmond WA, USA). The camera was installed in top-down perspective at a height of 345 cm above ground level. Each cow that entered the corridor passed a RFID-antenna (DeLaval AB, Tumba, Sweden). Cow identification triggered the recording of the video. The trigger signal of the photocell was captured and transferred to the computer by a low-cost USB digital input/output device (NI USB-6501, National Instruments, Austin Texas, USA). Depth recordings were made at 30 fps. The recording automatically stopped when a new cow was identified or if the photocell laser-beam of the RFID-unit was cut.
Each cow that passed the setup was identified by the RFID-antenna, and the timestamp of the identification was used to identify the individual cow in the recorded video.

2.3. Video processing: extracting image features

Video pre-processing filtered out the videos that contained (a) bad quality images; (b) multiple cows in the video; (c) not enough frames for analysis (d) an irregular cow gait (stop or run). For extracting animal based measures relevant for lameness detection, the full cow body (head to tail) needed to be segmented in the video. After video recording, cow identification merging and filtering, the remaining videos were further analysed. In these videos, the algorithm automatically segments the cow body in the images, extracts the back spine contour line, and calculates the back curvature of the cow. The back curvature of the cow is quantified by several feature variables: theta1, theta2, theta3, L-distance, Back Posture Measure (BPM) and inverse radius (See the work of Viazzi et al. (2014) and Van Hertem et al. (2014) for detailed description of the feature variables). The entire recording process ([a] trigger the video recording, [b] cow identification, [c] video pre-processing and [d] video analysis) was done automatically.
All recorded videos contained a depth recording and were saved to a 1 TB hard disk (Western Digital, Irvine California, USA) as .oni-files. The OpenNI 1.0 Software Development Kit framework was used to make the recordings with the Kinect camera.

2.4. Behaviour and performance data acquisition

The milking parlour was equipped with an electronic milk yield meter (FI2™, DeLaval AB, Tumba, Sweden) at every milking stand. During each milking, the individual milk yield was measured and expressed in kilograms. The daily milk yield was obtained as the sum of the milk yields of the two milking sessions in the day (morning, evening). In addition, milk conductivity and milk flow rate were also measured in each milking. Daily conductivity and flow rate values were obtained by averaging the values from the two milking sessions. Milking order was calculated from the milking time stamp that was registered in the management software.
All cows in the farm were equipped with a neck collar tag (DeLaval Activity Tag, DeLaval AB) used primarily for automated heat detection. Activity related to head movements was a filtered signal expressed by an index ranging from 0 to 255 bits per 1-h interval. The activity index was proportional to the number, amplitude and direction of the head movements. The activity data was split in two time frames: day time activity and night time activity. Day time was defined as the period of the day between 08.01 and 20.00, whereas night time was defined as the period between 20.01 and 08.00.
All data were transferred automatically during each milking to the herd management software (ALPRO™, DeLaval AB), and reports were extracted from this software in a spreadsheet format (Microsoft Excel, Microsoft Corp., Redmond WA, USA).

2.5. Reference: locomotion score

From 1 September 2013 until 30 April 2014, all cows in the milking herd were locomotion scored by two trained observers. Two observers were used to reduce the workload. In total 40 locomotion scoring sessions were done spread in time. Each time, the observer was positioned at the end of the corridor behind the spray box, and watched all cows pass in flank

view perspective. The locomotion score (LS) was based on the discrete 5-point numerical score of Sprecher et al. (1997) [1 = healthy; 5 = severely lame].
For model development, a binary reference is necessary. Therefore a cut-off threshold was used to divide the dataset in two different groups based on locomotion score. In a first instance, groups were divided as non-lame (LS = [1,2]) and lame (LS = [3,4,5]). In a second run, the severely lame cows (LS = [4,5]) were separated from the other cows (LS = [1,2,3]).

2.6. Lameness classification model

For further analysis, only complete (video-related feature variables, milk yield recordings, and daily activity measures) cow data were used. After elimination of incomplete data, the final dataset comprised of 3439 cow-observations.

2.6.1. Receiver Operating Characteristic (ROC)-curve

The ROC-curve plotted the sensitivity in function of the inspecificity (1 – specificity) of a binary classifier when varying its discrimination threshold. The sensitivity is a measure for the ability of the model in detecting true positive cases (in this case lameness). Specificity is a measure for the ability of the model in detecting true negative cases (in this case non-lame cows). The accuracy of the classifier was measured by the area under the ROC-curve (AUC). An area of 1 represented a perfect test; an area of 0.5 represented a worthless test. A diagnostic test can be classified as excellent (AUC = [0.9 – 1.0]), good (AUC = [0.8 - 0.9]), fair (AUC = [0.7 – 0.8]), poor (AUC = [0.6 – 0.7]) and fail (AUC = [0.5 – 0.6]), although the interpretation of ROC-curves are context specific (Bradley, 1997).

2.6.2. Forward stepwise binary logistic regression

A logistic regression model used a combination of *n* model input variables to fit the response variable *z* (Hosmer, & Lemeshow, 2000). For a binary logistic regression, the response variable *z* was dichotomous (lame vs non-lame). In model training, the model input variables were used to estimate the model parameters β_i in equation 1.

$$z = \beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \dots + \beta_n \times x_n. \qquad \text{(Equation 1)}$$

The *z*-value resulting from equation 1 was used to calculate the probability of the model output being a positive class member. In this case, a positive class member refers to lame animals, negative class members to non-lame animals. The probability was calculated in equation 2.

$$\text{Probability} = \exp(z) \,/\, [\,1 + \exp(z)] \qquad \text{(Equation 2)}$$

Stepwise regression includes models in which the choice of predictive input variables is carried out by an automatic procedure. In this study, a sequence of F-tests with significance level alpha set to 0.05 was used for this procedure.
Forward selection involves starting from no variables in the model, and tests the effect of adding each variable to the model using a chosen model comparison criterion. Only the variable (if any) that improved the model the most was then added to the model, and this process was added until no further improvement was possible. In the current study, the initial model was forced to start with including the BPM-input variable.
The MatLab Statistics Toolbox (MATLAB R2011b, the MathWorks Inc., Natick MA, USA) was used to estimate the model parameters in the logistic regression model.

3. Results

Table 1 shows the Area Under the ROC-curve (AUC) values for binary classification of lameness for each variable. The best discriminator for lameness is Theta2 (AUC = 0.7199, followed closely by BPM (AUC = 0.7021). Increasing the cut-off threshold for the locomotion score for binary transformation, increases the AUC-values on average by 0.0349 ± 0.0281. The highest AUC values were obtained for the Video variables (range [0.5452 – 0.7199]), followed by the activity related variables (range [0.5397 – 0.6155]) and the milk related variables ([0.5359 – 0.5846]).

(Insert Table 1 here)

The final multivariate model in the forward stepwise binary logistic regression procedure included 10 variables, which are when listed chronologically: BPM, day-time activity, theta1, walking speed, daily activity, milk conductivity, milk yield, milk peak flow rate, milking order and lactation stage (Table 2). The final AUC-value of the model was AUC = 0.7610, indicating an increase of 0.0411 compared to the highest obtained AUC-value with a single variable classifier.

(Insert Table 2 here)

4. Discussions

The results show that the combination of video data with behaviour and performance data improved the classification rate of a lameness classification model.

In the presented study, a logistic regression model was used to combine the computer vision data and the behaviour and performance data. The obtained accuracy rate (AUC = 0.7610) shows that the resulting model can be considered as a fair diagnostic test. Other types of models such as classification trees might improve the classification rate for lameness detection purposes.

The feature variables extracted from the video images reached higher accuracy values (AUC = [0.5452 – 0.7199]) than the activity- (AUC = [0.5397 – 0.6155]) and milk- (AUC = [0.5359 – 0.5846]) related variables. The work of Viazzi et al. (2014) and Van Hertem et al. (2014) showed a strong correlation between BPM and locomotion score. Therefore, BPM can be considered as a specific measure for lameness. Changes in behaviour and performance can be due to several diseases such as mastitis (Huybrechts, Mertens, De Baerdemaeker, De Ketelaere, & Saeys, 2014). Therefore, behavioural changes are unspecific for lameness. However a change in behaviour is often observed in lame animals, especially in acute cases (Reader, Green, Kaler, Mason, & Green, 2011; Van Hertem et al., 2013).

The study of Kamphuis et al. (2013) reported similar AUC-values for the univariate models, although different types of sensors were used and 14 day data windows were used for analysis. In this study, only data on the day of locomotion scoring were used. The combination of the computer vision data source, the activity sensor data source and the milking data source in to a multivariate model improved classification accuracy (AUC = 0.7610). Similar results were found by Kamphuis et al. (2013), where the combination of three sensor data sources into a multivariate model improved the classification accuracy to AUC = 0.74.

The behaviour and performance data were transformed to daily variables. Animal behaviour is influenced by a diurnal pattern. In this study, the activity was calculated for day- and night-time periods. The activity during day-time period (AUC = 0.6155) reached a higher accuracy than during the night time period (AUC = 0.5397). During night-time, the cow is most of the time resting, while during day-time, the farm routine (feeding, cleaning, etc.) might influence animal behaviour. The study of Van Hertem et al. (2013) also showed that a day and night separation of the activity and rumination data were more correlated to lameness than the daily values.

In the current study, a group level comparison between individuals was done. It is suggested by Viazzi et al. (2013) that changes from individual behaviour might improve lameness classification. Changes from normal behaviour and performance can only be detected when enough historical data are available. Future research should reveal if changes from individual behaviour are better indicators for lameness than group level statistics.

5. Conclusions

Univariate lameness classification performance was best for variables extracted from videos. This shows that the extracted feature variables are more specific than behaviour and performance variables. Lameness classification performance improved when the cut-off threshold for the locomotion score increased, indicating that severely lame cows are easier to distinguish from the herd.
The combination of specific video data with behaviour and performance data in a multivariate logistic regression model resulted in an improved lameness classification performance (AUC = 0.761) compared to the best univariate model (AUC = 0.7199).

6. Acknowledgements

This study was part of the EU Marie Curie Initial Training Network BioBusiness (FP7-PEOPLE-ITN-2008) and part of the Industrial Research Fund (IOFHB/13/0136), which the authors thank for their financial support.

7. References

Bradley, A.P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7), 1145-1159.

Bruijnis, M.R.N., Beerda, B., Hogeveen, H., & Stassen, E.N. (2012). Assessing the welfare impact of foot disorders in dairy cattle by a modeling approach. *Animal*, 6(6), 962-970.

Flower, F.C., & Weary, D.M. (2009). Gait assessment in dairy cattle. *Animal*, 3(1), 87-95.

Hosmer, D.W., & Lemeshow, S. (2000). *Applied Logistic Regression.* (2nd ed.). New York: John Wiley & Sons Inc., (Chapter 1-4).

Huybrechts, T., Mertens, K., De Baerdemaeker, J., De Ketelaere, B., & Saeys, W. (2014). Early warnings from automatic milk yield monitoring with online synergistic control. *Journal of Dairy Science*, 97(1), 1-11.

Ito, K., von Keyserlingk, M.A.G., LeBlanc, S.J., & Weary, D.M. (2010). Lying behavior as an indicator of lameness in dairy cows. *Journal of Dairy Science*, 93(8), 3553-3560.

Jonsson, R., Blanke, M., Poulsen, N.K., Caponetti, F., & Hojsgaard, S. (2011). Oestrus detection in dairy cows from activity and lying data using on-line individual models. *Computers and Electronics in Agriculture*, 76(1), 6-15.

Kamphuis, C., Frank, E., Burke, J.K., Verkerk, G.A., & Jago, J.G. (2013). Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological charachteristics of dairy cows to detect lameness. *Journal of Dairy Science*. 96(11), 7043-7053.

Kramer, E., Cavero, D., Stamer, E., & Krieter, J. (2009). Mastitis and lameness detection in dairy cows by application of fuzzy logic. *Livestock Science*, 125(1), 92-96.

Pluk, A., Bahr, C., Leroy, T., Poursaberi, A., Song, X., Vranken, E., et al. (2010). Evaluation of step overlap as an automatic measure in dairy cow locomotion. *Transactions of the ASABE*, 53(4), 1305-1312.

Pluk, A., Bahr, C., Poursaberi, A., Maertens, W., van Nuffel, A., & Berckmans, D. (2012). Automatic measurement of touch and release angles of the fetlock joint for lameness detection in dairy catlle using vision techinques. *Journal of Dairy Science*, 95(4), 1738-1748.

Rutten, C.J., Velthuis, A.G.J., Steeneveld, W., & Hogeveen, H. (2013). Invited review: Sensors to support health management on dairy farms. *Journal of Dairy Science*, 96(4), 1928-1952.

Reader, J.D., Green, M.J., Kaler, J., Mason, S.A., & Green, L.E. (2011). Effect of mobility score on milk yield and activity in dairy cattle. *Journal of Dairy Science*, 94(10), 5045-5052.

Sprecher, D.J., Hostetler, D.E., & Kaneene, J.B. (1997). A lameness scoring system that uses posture and gait to predict dairy cattle reproductive performance. *Theriogenology*, 47(6), 1179-1187.

Van Hertem, T., Maltz, E., Antler, A., Romanini, C.E.B., Viazzi, S., Bahr, C., et al. (2013). Lameness detection based on multivariate continuous sensing of milk yield, rumination and neck activity. *Journal of Dairy Science*, 96(7), 4286-4298.

Van Hertem, T., Viazzi, S., Steensels, M., Maltz, E., Antler, A., Alchanatis, V., et al. (2014). Automatic lameness detection based on consecutive 3D-video recordings. *Biosystems Engineering*, 119(1), 108-116.

Viazzi, S., Bahr, C., Schlageter Tello, A., Van Hertem, T., Romanini, C.E.B., Halachmi, I., et al. (2013). Analysis of individual classification of lameness using automatic back posture measurement in dairy cattle. *Journal of Dairy Science*, 96(1), 257-266.

Viazzi, S., Bahr, C., Van Hertem, T., Schlageter Tello, A., Romanini, C.E.B., Halachmi, I., et al. (2014). Comparison of a three-dimensional and two-dimensional camera system for automated measurement of back posture in dairy *cows. Computers and Electronics in Agriculture*, 100(1), 139-147.

Table 1: Tabular overview of the classification performance expressed as Area Under the ROC-Curve (AUC) of each variable in relation to the binary reference. Two different procedures are used to transform the 5-point numerical locomotion score to a binary level: LS = [1,2] as non-lame and LS = [3,4,5] as lame (column 3); LS = [1,2,3] as non-lame and LS = [4,5] as lame (column 4). The origin of each variable (Video, Activity and Milk) is stated in column 2. The rank of each variable is based on the values in column 3. For the analysis, 3439 cow-observations were used.

Variable	Variable class	AUC [1,2] vs [3,4,5]	AUC [1,2,3] vs [4,5]	Rank
Theta2	Video	0.7199	0.7868	1
Back Posture Measure	Video	0.7021	0.7756	2
Theta3	Video	0.6745	0.7499	3
Inverse radius	Video	0.6724	0.7468	4
L-distance	Video	0.6715	0.7456	5
Number of Frames	Video	0.5963	0.6233	7
Walking Speed	Video	0.5722	0.6038	11
Theta1	Video	0.5452	0.6063	13
Daytime activity	Activity	0.6155	0.6397	6
Daily activity	Activity	0.5898	0.6075	8
Night-time activity	Activity	0.5397	0.5452	15
Milk peak conductivity	Milk	0.5846	0.6033	9
Milk conductivity	Milk	0.5789	0.5963	10
Milking order	Milk	0.5560	0.5890	12
Milk peak flow rate	Milk	0.5444	0.5482	14
Daily milk yield	Milk	0.5372	0.5557	16
Lactation stage	Milk	0.5359	0.5270	17

Table 2: The forward stepwise binary logistic regression model included the following variables. Coefficients and standard errors of the coefficients are given in column 2 and 3. The order in which each variable was included in the model is given in column 4. The binary reference was constructed by dividing the dataset on two groups based on locomotion score: LS = [1,2] as non-lame, and LS = [3,4,5] as lame group.

Variable	Coefficient	Standard error	Step
Constant term	-15.8804	1.5034	0
Walking Speed	-3.4867	0.6163	4
Theta1	0.0658	0.0078	3
BPM	15.1437	0.8320	1
Daily activity	-0.0021	0.0002	5
Daytime activity	0.0014	0.0003	2
Daily milk yield	-0.0664	0.0142	7
Milk conductivity	0.2346	0.0835	6
Milking order	0.4257	0.1399	9
Lactation stage	-0.0009	0.0005	10
Milk peak flow rate	0.0996	0.0259	8