Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Biosystems and Technology

Cow Claws on Concrete Slats in the Waiting Area in a Dairy Barn Estimated by use of Image Analysis

H. Ardö[2], O. Guzhva[1], M. Nilsson[2], A. Herlin[1].

[1]*Swedish University of Agricultural Sciences, Department of Biosystems and Technology, Alnarp, Sweden.*
[2]*Lund University, Center for Mathematical Sciences, Lund, Sweden.*

Report from the Department of Biosystems and Technology, Swedish University of Agricultural Sciences

Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Biosystems and Technology

Abstract

Slatted concrete floors are commonly used in dairy barns for aisles, feeding and waiting areas. Maximum slot opening in Sweden is 35 mm with a maximum of 28% opening area for adult cattle in order to provide the adequate claw support. The construction of the slats has to consider this together with the length of the slats and the load from the weight of the animals on the slats. Presently, the dimension of the load strength of slats is based on assumptions and experience. An alternative approach is to estimate the true load of the animals on the slats by observation of animal distribution on slatted floors. The purpose of this study was to investigate possibilities of using machine learning algorithms and image analysis for assessing actual distribution of animals in the areas of interest and maximal weight load per slat element per unit of time. Images for the study were acquired from three surveillance cameras placed in the ceiling above the common waiting area (size 6x18 m) with entrances to four automatic milking systems (AMS). Then images were used to train a convolutional neural net classifier to detect and locate the cows in the images. Then, a probability distribution of where the claws might be located was constructed. By using this distribution in a Monte Carlo simulation, a probability distribution of the number of claws on each slat could be estimated, and from that, a worst-case estimate of the actual weight load was constructed. Results indicate that the 95% percentile number of claws on 160 mm wide slat area (slat width including the opening) was estimated to 3.03 and on a 560 mm slat area width was 5.63. Cows mounting was found in 7 of 9215 (0.2 %) examined pictures. The method proposed in this report was promising and for this purpose and could be used for practical assessment of animal distribution and loading and thus be a part of the dimensioning of construction.

Keywords: dairy barn flooring, deep learning, weight distribution, standards for concrete, precision livestock farming

Table of contents

1   Introduction

Slatted concrete floors are commonly used in dairy barns for aisles, feeding and waiting areas e.g. prior to milking. The design of slatted floors includes good drainage capacity obtained by the slot width or the void ratio (Magnusson et al., 2008). Maximum slot opening in Sweden is 35 mm with a maximum of 28% void ratio for cattle other than calves in order to give the cattle claw adequate support. The common slat width is 125 mm and with slot openings of 35 mm, gives a void area of 21.8%.

The construction of the slats has to consider this together with the length of the slats and the load from the weight of the animals on the slats in order to dimension the load strength. Presently, the calculation of the strength of the slats is set by a European standard which is entirely adopted by Sweden (SIS, 2007). The loads used in the calculations are based on the type and mass of animals and put into load classes. Three variable characteristics loads shall be taken: vertical characteristic linear and point loads and a horizontal characteristic point load. However, the calculation on load strength, considers a twin or a multiple slat construction instead of the prevalent single beams used in Sweden.

The use of different computer vision based systems for animal tracking and monitoring in dairy barns is rapidly developing area within Precision Livestock Farming (PLF). There are solutions suited for segmenting animals from the background in different areas of a dairy barn and distinguishing between different behavioural states (e.g. standing, laying down etc.) (Porto et al., 2015). Combining these methods with advanced machine learning approach (Simonyan and Zisserman, 2015), it is possible to create an algorithm capable of complex scene evaluation and multifactorial analysis of a dairy barn environment in relation to desired hypothesis.

The present study aims to estimate the presence of claws on individual slats by observation of animal distribution on the slatted floor in a waiting area to robotic milking. This would give more solid background data for determining the dimensioning of the strength of the slats. This study used the possibilities of machine learning algorithms and image analysis for assessing distribution of animals and claws in the area of interest and on the single slats.

2   Materials and Methods

2.1   Setup for recordings and initial image preparation

The video data for this study was acquired at the commercial free stall dairy barn equipped with four automatic milking stations (VMS®, DeLaval, Sweden) and a joint waiting area to these stations, that was used as the recording site. The size of the waiting area was 5.3x16 meters and 252 Swedish Holstein cows had free access to all four automatic milking stations (AMS) at any time during the day. The floor at the waiting area consisted of the concrete slats with 125 mm in width and the slat opening of 35 mm. In order to acquire videos, setup with three Axis M3006-V (Axis® Communications) cameras placed at 3.6 meter height and a Hewlett Packard ProLiant Microserver for data storage was used. All the acquired videos were recorded with the frame resolution 800x600 pixels, RGB colour space and a framerate of 16 frames per second.

In order to create the pool of training data for the classifier, random images (n=1722) were taken from the recorded videos (7TB of video material) to form a representative sample for different weekdays and hours. The landmark points (head, left and right shoulder, front middle, left and right hip and back middle) were then manually assigned to every cow present in an image. The information from every landmark point, containing image coordinates and absence or presence of occlusion, was stored in .JSON files, named to match the unique image code with camera number, date and time stamps.
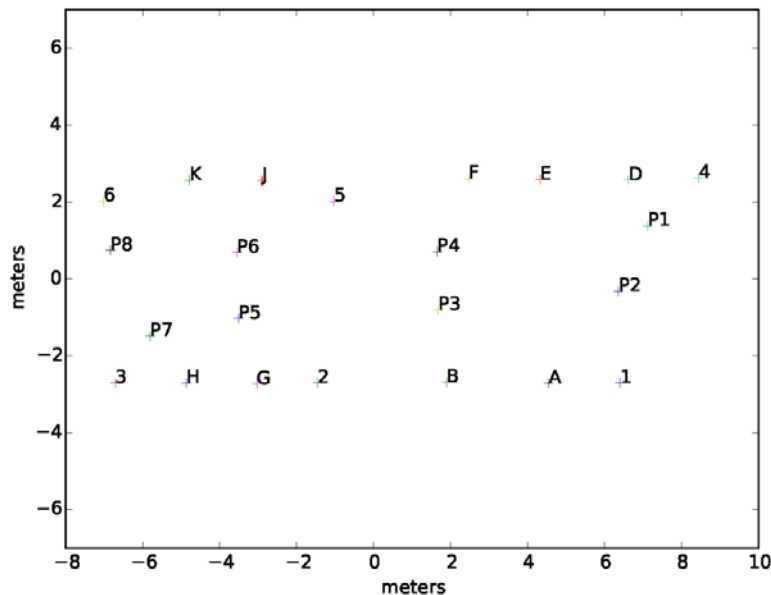
The use of landmark points for shoulders and hips, instead of actual positions of claws, was implemented, as the claws are often not visible in the image as they are occluded either by the cow itself or by another cow. These points are located above the claws, but all the cows in this farm are of approximately the same height. A virtual slated floor could be placed at this height in the scene to allow the analysis to be performed using these points instead of the actual claws.

2.2   Calibration of cameras and plane estimation of the study area

In order to fulfil the specific research questions regarding dairy barn slatted floor and distribution of animals on it, correct transition of image coordinates into real-world coordinates was crucial. To assure that all the coordinates can be reliable and that the observer/algorithm will be able to correctly identify all the objects, three images were merged and synchronized after applying normalization algorithms. Setups with only one camera (even with wide observation angle) could suffer from a number of image artefacts (e.g. radial

distortion, tangential distortion, occlusion between objects/cows) therefore; it was decided to use three cameras for a relatively small region of interest (ROI) in the waiting area.

For this study, the classical pinhole camera model augmented with a lens distortion model was used for scene view reconstruction (Tsai, 1986). The scene view was formed through the projection of 3D world points into the image plane and to assure correct disposition and perspective of objects in a merged image covering the ROI, we used number of planar markers to estimate scene homography and lens distortion. The camera calibration method developed by Tsai (1986) and further improved by Horn (2000) includes both interior and exterior orientations, corrections for the distortion and a scale factor for the reliable correlation between target and scene coordinates. The number of planar markers was used for the camera calibration on site (Figure 1). These markers were placed at the height of the virtual floor. This height was estimated to be 1.49 meters with a standard deviation of 0.05 by measuring twelve random cows in the study area.

Figure 1. The location of the calibration markers on the virtual floor 1.49 meters above the real floor used to calibrate the cameras. The crosses indicated marker positions and next to the cross is a label used to identify a specific point.

Once the calibration is done, the virtual floor can be projected into the image as shown in Figure 2 below. The floor is represented with a blue grid with one grid element for each 160 mm slat+slot. The annotated cows are shown as red H-shapes where the endpoints indicate the positions of the claws. It is also possible to project the annotated cows from the images into the coordinate system of the virtual floor, as is show in the second row of the Figure 2.
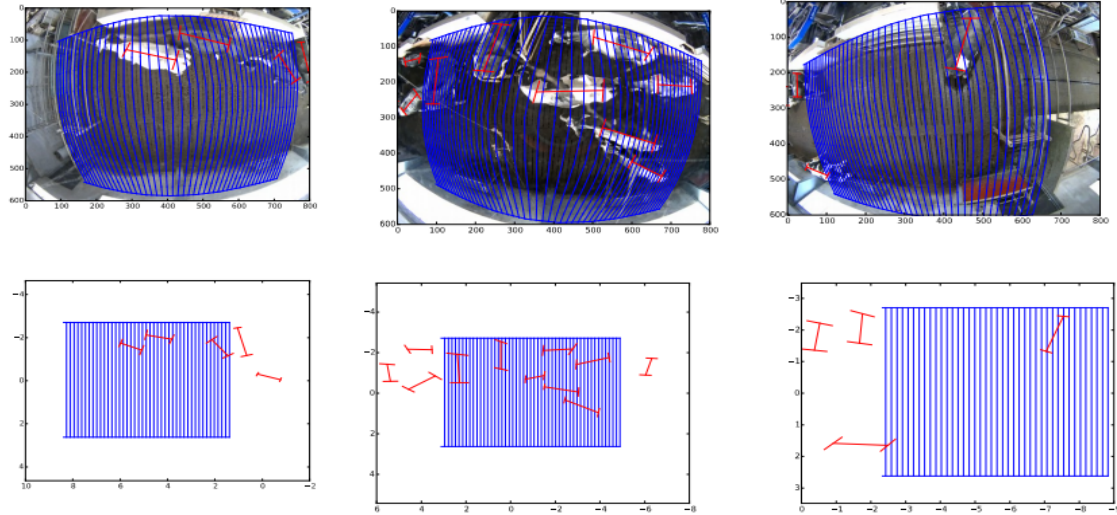
Figure 2. Projection of the virtual floor over barn floor.

2.3 Convolutional neural network (CNN)-based cow detector, claw and slatted floor models

The images produced by the cameras were de-warped to compensate for the lens distortion and rotated to form images that are distortion-free with a camera orientation perpendicular to the ground plane. In such a setup the image of a single cow does not vary in size as the cow moves around in the image. A deep convolutional neural network was trained on those input images to detect four points on each cow (head, neck, centre and rear). The output of this network is a multichannel image with probabilities for the different parts as well as for the ground. Using those probabilities as an input a second network was trained to detect the centres of the cows and their orientations. The orientation is parameterized as a discretized angle, where a full circle is divided into 32 different orientations. The second network's output consists of 33 different classes, one for each orientation and one for the ground (i.e. points in the image where there is no cow). The figure below shows some example results. The three de-warped images from the cameras are stitched together and the probabilities of the different parts of the cows are shown in different colours. The final detections (position and orientations) are shown as yellow lines (Figure 3).
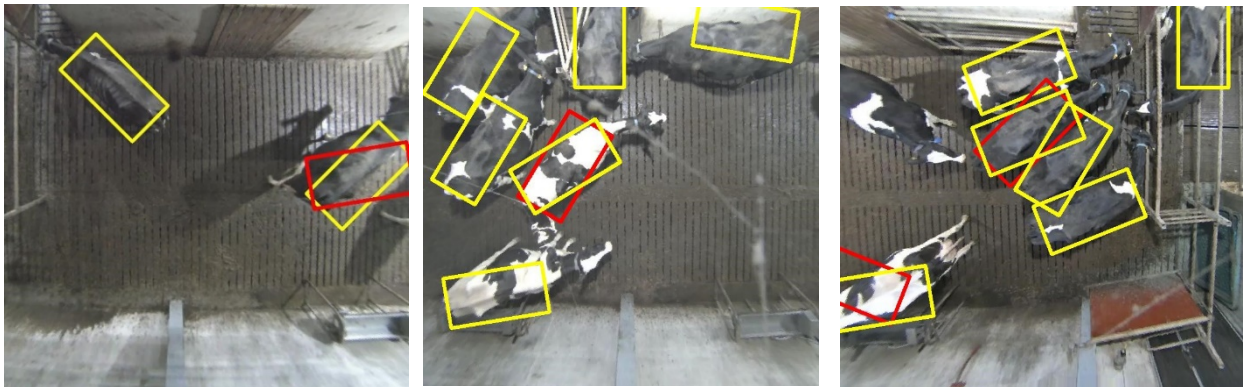

Figure 3. Example of output image from CNN with colour zones detected and assigned to every cow in a scene (Magenta is head, red is front middle, green is centre and blue is rear middle).

8

The network used to detect the cow parts consists of 11 weight layers and is based on VGG-11 (Simonyan and Zisserman, 2015). The first layer has 32 feature-outputs and then the number of features is doubled after each maxpool layer. There are no fully connected layers at the end. Instead, 1x1 convolutional layers are used. This means that we can present an entire image to the network and the entire image will be scanned in the same fashion as sliding window detector works. To increase the resolution of the resulting part probability images, four versions of the input image translated horizontally and vertically and fed to the network.

The networked was trained using stochastic gradient descent with momentum (Rumelhart et al., 1986). An initial learning rate of 1.0 was used and it was lowered by a factor 1/10 each time the validation loss flattened out. A batch size of 256 and a momentum of 0.9 was used. The network was regularized using weight decay of 0.0001 and batch normalisation (Ioffe and Szegedy, 2015). The training data consisted of 1722 images, where seven anatomically related points were marked on each visible cow (head, front middle, left and right shoulder, left and right hip and rear middle). In total 6399 cows and 44793 points were annotated. In addition to that, one synthetically centred point for every cow was generated as the mean between the front middle and the rear middle. All the images were divided into a training set (90 %) used to fit the network parameters and one validation set (10 %) used to monitor its generalization performance.

After the network produced the detections, overlapping detections were removed by a pruning state. The Figure 4 below shows some results from the detector where the detected cows are marked with rectangles and the detections removed by the pruning are marked in red.

Figure 4. Example of the output image from the cow-detector

The cow detector only generates a centre position and orientation of each cow. To estimate the position of the claws from such detections, a statistical model was formed from the manual annotations. All the annotated cows were normalized by translating their centre to (0,0) and rotating them to align their body with the x-axis. The normalized positions for all four claws were then plotted in different colours in the Figure 5 below, and a mean cow shape was estimated by taking the mean position of each claw. This mean shape was plotted as a red H-like shape with endpoints of the lines marking the claw positions.
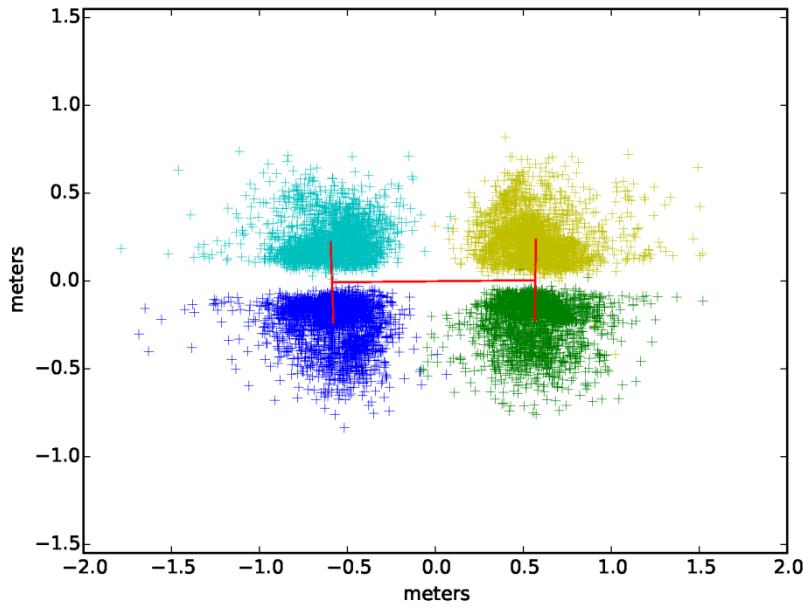
Figure 5. Deviation of manually annotated claw coordinates (endpoints of red H shape figure) compared to automatic detection for all four claws.

To simulate a floor with 125 mm wide slats and slot openings of 35 mm, a grid with 160 mm wide rectangles was placed in the image (slat area). The rationality here is that if a claw was placed over the opening the full weight supported by that claw will still be placed on the slat. Accordingly, a slat area rectangle with 560 mm width was simulated in order to represent a construction of multiple slats. The number of claws placed in each rectangle were then calculated and divided by the total number of observed images. This gives a probability distribution over the number of claws on a random slat at a random point in time. This analysis was performed both using the manual annotated cows (1722 images) and using automated detections on a separate set of images (5861 images) that was not used during the training of the detector. For each of the cows detected by the CNN detector, a random sample was drawn from the claw distribution presented above. This sample was then translated and rotated to be placed at the detected centre at the detected orientation. By comparing those results, the precision of the automated process could be estimated.

From the probability of there being exactly n claws on a random slat at a random time, p(X=n), the probability of there being n or less claws can be estimating as a sum from 0 to n:

$$p(X \leq n) = \sum_{i=0}^{n} p(X = i)$$

The probability distribution of the worst case at any random point in time is found by taking the maximum over all slats. On a floor with m slats, the probability of there being n claws on the maximum is that same as there being n on all the slats. If the slats are assumed to be independent this can be estimated as:

$$p(\max(X) \leq n) = p(X \leq n)^m$$

2.4   Cows mounting

A total of 9215 images were randomly sampled from a random camera at a random time were selected from the video dataset used above. Every picture was manually classified on the presence or non-presence of cows mounting each other.

3   Results and Discussion

3.1   Claws on slats

The first study compared 160 mm wide slats with 560 mm wide slats. The area studied was 16 meter wide, which means that 100 slats were needed to cover the entire area if they were 160 mm wide, while 29 was enough for the 560 mm case. The slats were assumed to cover the entire height of the area (5.3 meters). Experiments were performed both with manually annotated cows and with automatically detected cows. The manual cases consisted of 1722 images randomly sampled from a random camera at a random time. This resulted in a total of 14930 claws and 77847 slat observations for the 160 mm case and 23079 observations for the 560 mm case. The automated detections were performed on 5861 images resulting in 508960 claws and 2631490 (23079) slat observations for the 160 (560) mm case. Results are shown in the Figure 6 below, in form of the probability that a random slat at a random time is loaded with n claws, as well as the probability that the maximum load on any single slat at a random time is n claws. The results indicate that the 95% percentile number of claws at a random time on 160 mm wide slat area was estimated to 0.84 and on a 560 mm slat area width was 2.55.
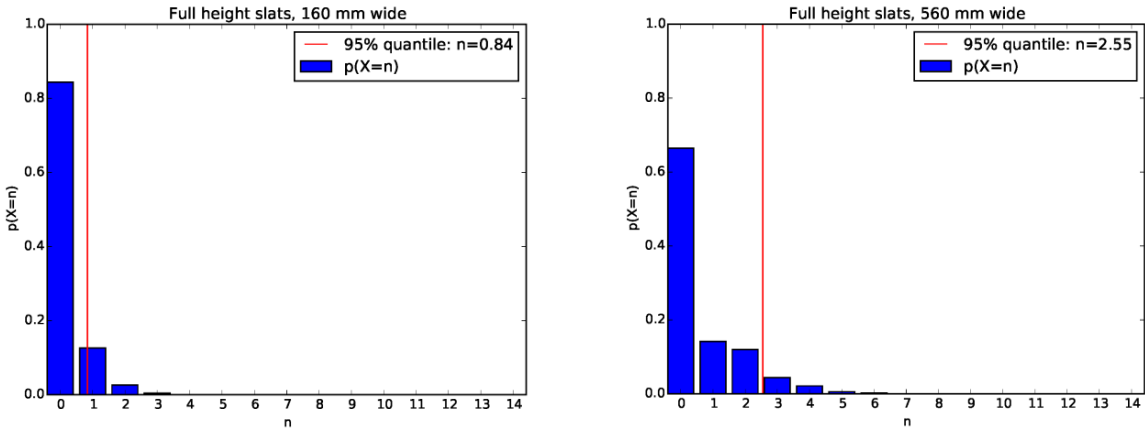

Figure 6. The probability of n claws placed on a random 160 mm wide slats area (left figure) and 560 mm wide slats area (right figure) at a random point in time. The 95% quantile of the distribution is estimated by interpolating linearly between the integer values of n and marked as a red line.

12

The maximum loading at the 95% percentile number of claws at a random time on 160 mm slat area was estimated to 3.03 and on a 560 mm slat area width was 5,63 (Figure 7).
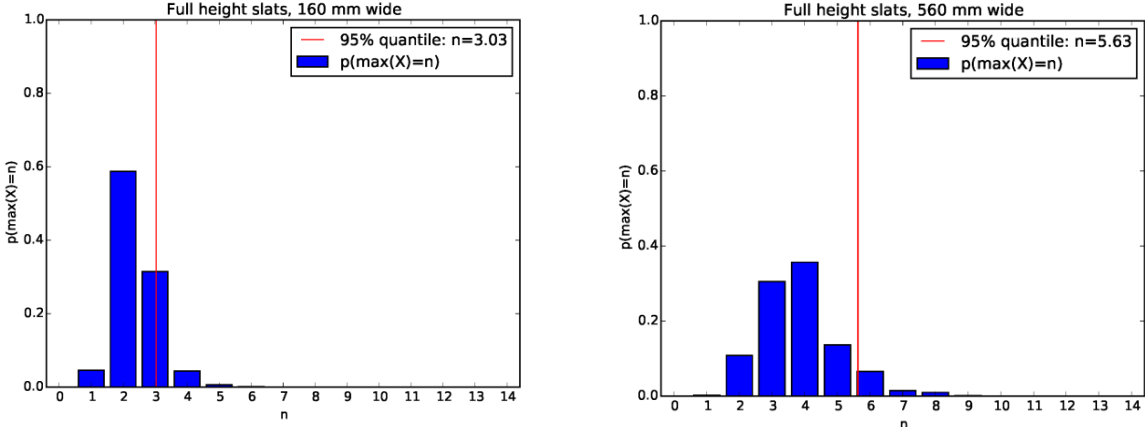

Figure 7. The probability of n claws placed on the most loaded slat in the waiting area on a 160 mm wide slat area (left figure) and on a 560 mm wide slats area (right figure) at a random point in time. The 95% quantile of the distribution is estimated by interpolating linearly between the integer values of n and marked as a red line.

The automated results correspond very well with the manual versions and the thinner 160 mm slats receive significantly less claws as would be expected. The automated approach slightly underestimates the probabilities as compared to the manual approach. This is probably due to the fact that the detector sometimes can fail to detect a cow.

3.2 Cows mounting

A total of 7 images with cows mounting was detected out of the 9215 examined images which is 0.2 %, representing about 4 minutes per day (Figure 8).
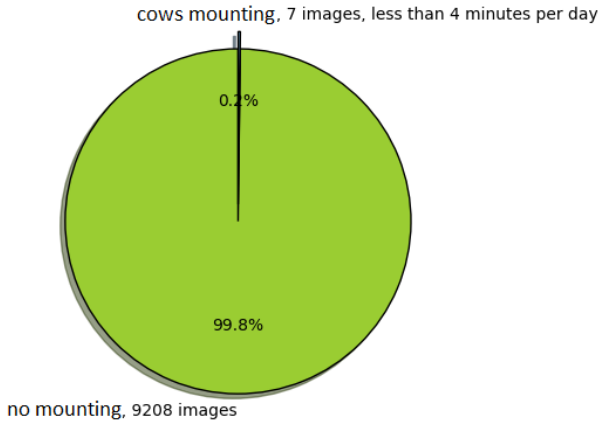

Figure 8. Observations of cows mounting and no mounting in 9215 images

4   Conclusions

The detector performance and proposed method for analysis showed the potential for further development and could be used as a tool for practical assessment of actual weight load/animal distribution in areas of interest. As expected, there was a large difference in the loading of claws on the 160 mm wide slat area vs. the 560 mm slat area. We therefore recommend that the calculations of the loading strength take this into account.

5   Acknowledgements

References

Horn, B.J.P., 2000. Tsai's camera calibration method revisited, http://people.csail.mit.edu/bkph/articles/Tsai_Revisited.pdf [2016 06. 18]

Ioffe, S. and Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift., arXiv:1502.03167v3

Magnusson, M., Ventorp M., Nilsson C., 2008. Drainage Capacity of Concrete Slatted Floors for Dairy Cattle. CIGR Journal Volume X. Available: http://www.cigrjournal.org/index.php/Ejounral/article/view/1029/1022# [2016-05-06]

Porto, S.M.C., Arcidacono, C., Anguzzo, U., Cascone G., 2015. The automatic detection of dairy cow feeding and standing behaviours in free-stall barns by a computer vision-based system. Biosystems Engineering (133): 46-55., doi:10.1016/j.biosystemseng.2015.02.012

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323 (6088): 533–536., doi:10.1038/323533a0.

Simonyan, K. and Zisserman, A., 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. Published as a conference paper at ICLR 2015, arXiv:1409.1556.

SIS, Swedish Standards institute. 2007. Precast concrete products – Floor slats for livestock. Stockholm, SIS Förlag AB (SS-EN 12737:2004+A1:2007).

Tsai, R.Y., 1986. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, USA, 364-374.