1 Theoretical foundation

The main hypothesis in this works involves investigating the possibility to extract complex cow behaviour from recorded video data for further semi-automatic/automatic annotations and analysis.

The hypothesis is confirmed by experimental validation with a newly developed method for detection and assessment of cow behaviour, based on 7-point shape-model [1] from our previous work. Combining these two principles, it would be possible to build the framework for a multi-component analysis of video fragments containing cow behaviour (sorting video fragments, registering and assessing social interactions, performing social network analysis based on proximity distances between individuals etc.).

The general theoretical technical detection parts are based on classical deep learning and convolutional neural networks extended to produce rotated rectangles as detections.

Animal behaviour and welfare can be studied/assessed by looking at different interactions occurring between the animals. Video recordings of a scene of interest are often made and then watched/evaluated by experts. Some recent studies on cow behaviour in the dairy barn environment were based on different GPS or wireless sensor network (WSN) solutions [2]. These solutions allow scientists to see spatial distribution of animals and to measure varying levels of activity [2]. However, the number and complexity of behaviours that could be monitored with the position-based approach is limited.Therefore, new methods based on video surveillance and image analysis, which could extend the number of parameters for studying, are of great importance. This is achieved using a cow detector that detects cows in the scene and represents them with rotated bounding boxes. That means the number of cows could be counted and distances between the cows can be measured. Rotated bounding boxes allow much more precise distance measures between diagonally oriented cows as compared to more classical non-rotated bounding boxes due to the elongated aspect ratio of the cows.

Most previous work on detecting cows by using video cameras have been focused on monitoring areas where the orientation of the cows was known due to physical limitations imposed by the surroundings. These limitations also restrict any complex social interactions between individuals within the group. Studying those interactions helps in advancing the understanding of driving forces behind all the day-to-day activities performed by animals, with high levels of correlation between positive/negative interactions occurred and health/welfare animal-based indicators. Therefore, given a tool for automatic/semi-automatic video data management (sorting away uninteresting fragments) and possibility to register/assess different behaviours/interactions is of great importance.

2 Main contributions

A novel approach towards cow interaction detection involving a two step CNN for rotated cow detection with a following interaction watchdog fitering out uniteresting video that is unlikely to contain interactions.

3 Closest work

The current state of the art for detecting cows freely moving around was presented by Porto et al. [3]. They use a Viola-Jones based detector and need 6 cameras at 4.6 meters height to cover a 15.4×3.8 m area and detects cows in 3 different orientations: vertical, horizontal and diagonal with a hit rate of 90%. They do not use separate detectors for the two different diagonals which mean that producing rotated bounding box from their results would not be straight forward. In the work presented in this paper, only 3 cameras at 3.6 meters height were needed to cover a 18×6 m area and detect cows in 32 different orientations with a hit rate of 97%. Note that the hit rates are from different datasets, and that the dataset used in this article have larger variations in viewpoints due to the use of fewer cameras at a lower height to cover a larger area.

4 State of the art

Section "1.1 Related work" contains a review of the state of the art and includes comparative evaluation of our work against the closes work mentioned above.

5 Originality

This submission is our original work. An earlier version was presented at the Visual observation and analysis of Vertebrate And Insect Behavior 2016 workshop in conjunction with 23rd International Conference on Pattern Recognition (ICPR 2016). A copy of that version is included in the submission as vaibardo.pdf Sections 1.1, 2, 5.2 and 5.3 are completely new while Section 1, 3 and 6 have been extened.

References

[1] O Guzhva, H Ardö, A Herlin, M Nilsson, K Åström, and C Bergsten. Feasibility study for the implementation of an automatic system for the detection of social interactions in the waiting area of automatic milking stations by using a video surveillance system. *Computers and Electronics in Agriculture*, 127:506–509, 2016.

[2] E. S. Nadimi, R. N. Jörgensen, V. Blanes-Vidal, and S. Christensen. Monitoring and classifying animal behavior using ZigBee-based mobile ad hoc wireless sensor networks and artificial neural networks. *Comput. Electron. Agric*, 82, 2012.

[3] Simona MC Porto, Claudia Arcidiacono, Umberto Anguzza, and Giovanni Cascone. 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, 2015.