Multimedia Tools and Applications (2020) 79:14711–14724
https://doi.org/10.1007/s11042-019-7344-7

Individual identification of dairy cows based on convolutional neural networks

Weizheng Shen[1] · Hengqi Hu[1] · Baisheng Dai[1,2] · Xiaoli Wei[1] · Jian Sun[1] · Li Jiang[1,3] · Yukun Sun[4]

Received: 30 August 2018 / Revised: 23 January 2019 / Accepted: 5 February 2019 /
Published online: 13 February 2019

Abstract
Individual identification of each cow is significant for precision livestock farming. In this paper, we propose a novel contactless cow identification method based on convolutional neural networks. We first collected a set of side-view images of dairy cows, then employed the YOLO model to detect the cow object in the side-view image, and finally fine-tuned a convolutional neural network model to classify each individual cow. In our experiments, a total of 105 side-view images of cows were collected, and the proposed method achieved an accuracy of 96.65% in cow identification, which outperformed existing experiments. Experimental results demonstrate the effectiveness of the proposed method for cow identification and the potential for our method to be applied to other livestock.

Keywords Cow identification · Precision livestock farming · Computer vision · Convolutional neural networks · Object detection

1 Introduction

For accurate management of farms, individual identification of dairy cows is one of the most important pieces of information, which can be used in automatic behavior analysis, milking, health monitoring, weighing, and other activities [12]. In addition, highly accurate cow

✉ Baisheng Dai
bsdai@neau.edu.cn

1 School of Electrical Engineering and Information, Northeast Agricultural University, Harbin 150030, China

2 Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling 712100, China

3 School of Electrical Engineering and Information, Heilongjiang Bayi Agricultural University, Daqing 163319, China

4 School of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China

identification can solve problems related to missing cows, swapped cows, and false insurance claims. Currently, many farms use ear tags with radio-frequency identification (RFID) sensors [7]. However, ear tags are intrusive and tend to get lost over time. Furthermore, an ear tag's reading/writing distance is limited [25]. Therefore, a contactless-based approach for individual identification of cows would be an improvement.

The use of computer vision techniques has been expanded significantly in recent years [15–17, 22, 27]. Researchers have made significant contributions to understanding image sequences, including some typical topics such as object detection [2, 3, 19–21] and semantic segmentation and labeling [1, 13, 26]. Computer vision techniques have also been widely used in animal biometrics and wildlife science [10]. By using computer vision, contact is not required for cows to be identified and their behaviors to be analyzed [8]. These noncontact methods have improved the practicality and automation of identification and reduced its cost.

In general, there are currently four areas used for cow identification: muzzle, face, back and trunk, as shown in Fig. 1(a)-(d). Kumar et al. [12] proposed a cow identification system using the muzzle area. Convolutional neural networks and deep confidence networks were used in the system to extract features of the cow's muzzle for cow identification. Although the system achieved a good accuracy, it was difficult to capture the images of the cow's muzzle since the muzzle area is so small. Kumar et al. [11] also presented a cattle identification method using facial images. This method first used the AdaBoost face detection algorithm [24] to detect cattle faces. Then, by applying PCA, LDA and ICA technology, the cattle face features were extracted. Lastly, support vector machine (SVM) and incremental support vector machine (ISVM) were trained to classify cattle faces. However, the unconstrained environment of head movements and uneven lighting conditions introduced complexity, which deteriorated the performance of the identification. Zin et al. [28] applied interframe difference and a horizontal histogram to extract the back images of cows placed on a rotary milking parlor and trained a convolutional neural network to identify individual cows. This method demonstrated that the pattern on a cow's back can be used to identify individual cows. Zhao and He [8] proposed a method for detecting a cow's trunk and then using convolutional neural networks to identify cow trunk images. The method used the interframe difference method to obtain a rough outline of the cow and analyzed the binary image by segmentation span analysis to locate the cow's trunk area. In this method, deep features were only extracted from the cow's trunk area. However, identification using the cow's trunk area depends not only on the object detection but also on the detection of the trunk area. Furthermore, beyond the trunk area, the cow object (Fig. 1(e)) in side-view images contains the head and legs. This method ignored the information of the head and legs, which also includes contour and texture features. Compared to the cow's muzzle and face, the side of a cow is more stable and easier to detect. In this paper, we verified the contribution of each area of the side of the cow (including head, legs, trunk and cow object) to individual identification. We also verified that cow identification can be performed directly using the cow object with good accuracy, without extracting the trunk area.

(a) (b) (c) (d) (e)

Fig. 1 Some areas of interest for cow identification: (**a**) muzzle, (**b**) face, (**c**) back, (**d**) trunk and (**e**) overall cow object

In this paper, to identify individual cows, we exploited all the information of the cow object in side-view images, including head, trunk and leg areas, by means of convolutional neural networks. First, we captured the side-view image sequences of the cows walking through the straight milking channel, then employed the YOLO [20] object detection model to localize the cow object in each frame of the sequences, and finally identified each of the cows using a fine-tuned convolutional neural network model.

The rest of this paper is organized as follows. In Section 2, we present the experimental materials of our work. The detailed proposed method is explained in Section 3, followed by the experimental results, analysis and discussion in Section 4. Finally, in Section 5, the conclusion is presented.

2 Materials

2.1 Image acquisition

The experimental image sequences were captured at Kangda Cow Farm, Shuangcheng, Harbin, China, in November 2017. The experimental object was the Holstein cow. After being milked, each cow passed through a milking channel in turn. An ASUS Xtion2 camera was placed next to the milking channel. The size of the field of view was specially set to a value greater than the size of each cow. A total of 105 cows' side-view image sequences were captured in our experiment. The image sequences were saved in .oni format with a resolution of 640 (horizontal) × 480 (vertical) pixels.

2.2 Experimental data

The 105 cows were numbered from 0 to 104, which was simpler than using their ear tag numbers. The duration of each cow in the milking channel was approximately 6 s, and during this time the whole cow was in the field of vision for approximately 2 s. The image sequences were sampled at 30 FPS. During the capture process, there were some cases of frame loss and adjacent cows overlapping. The captured image sequences were manually postprocessed to remove those overlapping frames of cows for the identification task of this work. The rest of the frames were arranged into corresponding cow categories to form the image set, which were prepared to be detected and identified. After using the YOLO object detection model to localize the cow objects in the original images, we obtained 1433 side-view images of cow objects. The side-view images of each cow were randomly divided into a training set and a validation set with a ratio of 70% and 30% respectively. We obtained 1015 images of the training set and 418 images of the validation set. The training set was specially augmented by random crop and horizontal flip to avoid the problem of overfitting. Finally, there were 82,215 images in the training set and 418 images in the validation set.

3 Method

The procedure of the proposed cow identification method is shown in Fig. 2. To suppress background noise and the interference of different cows, the input images were first sent into the YOLO object detection model for localizing each of the cow objects of interest, then the images of detected cow objects were fed into a fine-tuned convolutional neural network for the final identification. The details of the proposed method are described as follows.

3.1 Object detection

To detect the location of the cows in each image, the YOLO model was employed in this work, which is a deep learning framework for object detection, pretrained on MSCOCO [14]. This framework achieved real-time object detection rates and met the real-time needs of the individual identification of cows. The object detection results are shown in Fig. 3. We arranged the cows of interest for each image into the images to be identified with corresponding cow categories. The set of the images to be identified was selected by

$$COW = \{det_i | A_{det_i} > A_t, label_{det_i} = cow; i = 0, 1, \ldots n\} \tag{1}$$

where COW denotes the set of the images to be identified, A_{det_i} denotes the area of object detection result det_i, and A_t denotes a defined threshold value of area. The value of A_t is empirically set to $0.25 \times (640 * 480)$, with which the salient cows of interest in each image can be selected out. $label_{det_i}$ denotes the name of the object detection result det_i, and n denotes the number of object detection results of the image sequences. We selected out each object detection result whose object name was cow and area was greater than A_t for the following identification.

3.2 Individual identification

Recently, convolutional neural networks (CNNs) have achieved great success in visual recognition/classification tasks by learning the deep features of the input images [4–6, 9, 18, 23]. Considering that some cows have less texture information, it is difficult to manually define discriminable features for those cows with traditional algorithms. Therefore, in this paper, the convolutional neural network is used for identification and we propose a cow identification method based on a fine-tuned AlexNet model [9]. An overview of the proposed model is shown in Fig. 4. AlexNet is a convolutional neural network that has a strong classification performance in ImageNet competition. Here, we used the pretrained weights as the initial parameters of the model to fine-tune a new model with the abovementioned cow images for the purpose of identification. We extracted the deep features of the images by fine-tuning AlexNet and updating the parameters of the last two layers of AlexNet.

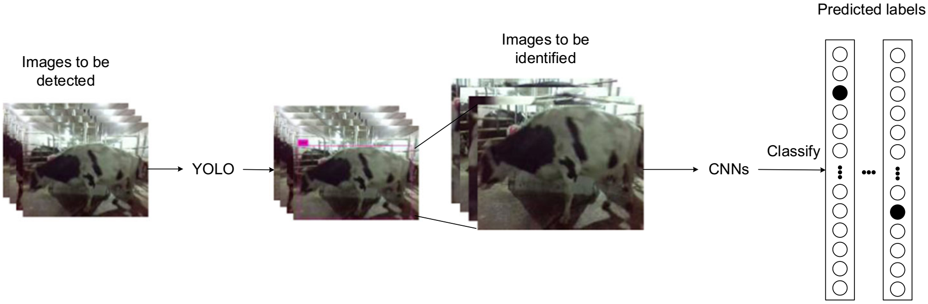

Fig. 2 Procedure of the proposed cow identification method

Fig. 3 Illustrations of cow object detection results

Specifically, the original images in the dataset were input into the network and multiple convolution and pooling operations were performed. Then, the final prediction result was generated through three consecutive fully connected layers. In this study, the number of cows was 105, so the output layer was stacked by 105 neurons, which represents the probability of classifying the input cow to a certain category.

After calculating the error between the prediction and the actual category number, the stochastic gradient descent method was used to minimize the loss function and realize the update of the weights. When the model converged, a set of optimal weights of the network was obtained. That was the model obtained by training.

Finally, in the evaluation stage, the cow images in the validation set were fed into the network and predicted by the fine-tuned model. At this stage, only forward propagation was performed, and the prediction category with the highest probability was the final identification result.

4 Experimental results and analysis

4.1 Identification results and analysis

For the experimental parameters, the batch size was 122, the initial learning rate was 0.01, and it was set by step to 1/10 of the previous in which the model loss function fluctuated slightly. Four hundred and eighteen images in the validation set were used to evaluate the model in each iteration. After 63 iterations, the accuracy reached 96.65%. The identification performance comparison of the proposed method and the existing works is shown in Table 1.

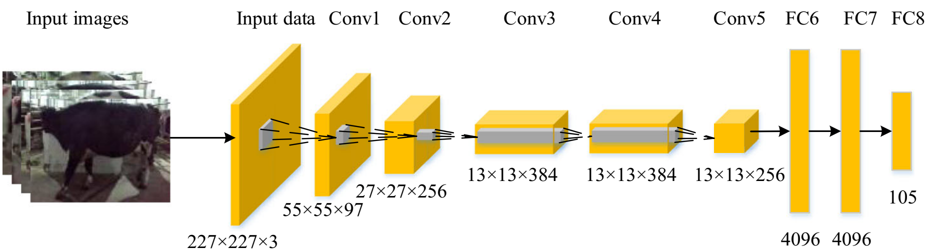

Fig. 4 Structure of AlexNet

The first two methods—image comparing and SIFT—are used in Zhao and He [8] to compare them with the main method (the third method) of their paper.

As seen in Table 1, since image comparing based on the pixel gray-values are sensitive to image deformation and light, the accuracy of individual identification of cows with this method was low.

In the work of Zhao and He [8], SIFT was used for individual identification as a comparative experiment. They extracted the trunk images of cows and used SIFT to match them, but the accuracy was low, and the experiment was time-consuming. We also used SIFT to identify the cow objects in our preliminary experiments. By analyzing the feature detection maps and corresponding feature matching of images, especially the cows with less texture information, we found that the features were mainly distributed in the background area of the images, and the matched features were mostly in this area, so the method is not reliable.

The deep features, such as the contours and textures of cows, can be extracted in convolutional neural networks, so Zhao and He [8] used a method based on convolutional neural networks to classify trunk images of cows. The method showed a much better performance, improving by more than 50 percentage points compared to image comparing and SIFT.

However, when the trunk area has no texture or less texture, the information on the head, legs and the overall contour of the cow cannot be ignored. Therefore, the proposed method achieved an accuracy of 96.65% for cow identification of 105 different cows and achieved an impressive 6.1% improvement in performance compared to the work of [8].

From the experimental results given above, it can be concluded that identification using cow object is better than using trunk area.

We also compared the performance of four different CNN models to identify dairy cows, i.e., VGG-16 [23], ResNet-50 [4], DenseNet-121 [5] and AlexNet. The identification performance of each model is listed in Table 2.

As seen from Table 2, the performance of AlexNet is superior to other models. It achieved 21.29%, 11.24% and 1.2% outperformance over DenseNet-121, ResNet-50 and VGG-16 respectively. One possible reason is that the amount and diversity of our experimental data was limited, due to the restriction of data collection on a business farm, and the model with high capacity was prone to the problem of overfitting.

4.2 Feature maps and analysis

Fig. 5 shows the feature maps output by the first convolution layer of our convolutional neural network.

From the 96 feature maps, it can be seen that the model effectively extracted the features, such as contours and textures of cows. The convolution operation had good filtering and smoothing effects, which can enhance signal characteristics and reduce noise. At the same

Table 1 Identification performance comparison of different methods

Method	Accuracy (%)	Object categories
Image comparing [8]	32.95	30
SIFT [8]	37.89	30
Zhao et al. [8]	90.55	30
The proposed method	96.65	105

Table 2 Identification performance comparison of four base networks

Base network	Accuracy (%)	Object categories
VGG-16	95.45	105
ResNet-50	85.41	105
DenseNet-121	75.36	105
AlexNet (the base network of proposed method)	96.65	105

time, the convolution operation can map the input image into different gray spaces, which ensures the network's ability to resist external interference under different lighting conditions.

Fig. 6 shows the three overlaying images of one input image and its corresponding feature maps of the last convolution layer in our convolutional neural network. From this figure, we found that the head, trunk and legs were obviously activated. Therefore, it can be concluded that all the information of side-view images, including head, trunk and leg areas, can be used to learn discriminable features for cow identification by means of convolutional neural networks.

4.3 Part contribution analysis

To analyze the contribution of each area to the individual identification of cows, we segmented the side-view images of cows into different parts, i.e., the head, the trunk and the legs, as shown in Fig. 7. By analyzing the geometrical position of the three areas in the side-view images, we first located the cow's trunk according to the work of Zhao and He [8] and then obtained the head and leg areas based on the geometric relationships.

Among the 105 cows, 12 cow head areas were obscured by the former cow while passing through the milking channel. To avoid sample variability, we used the remaining 93 cows in

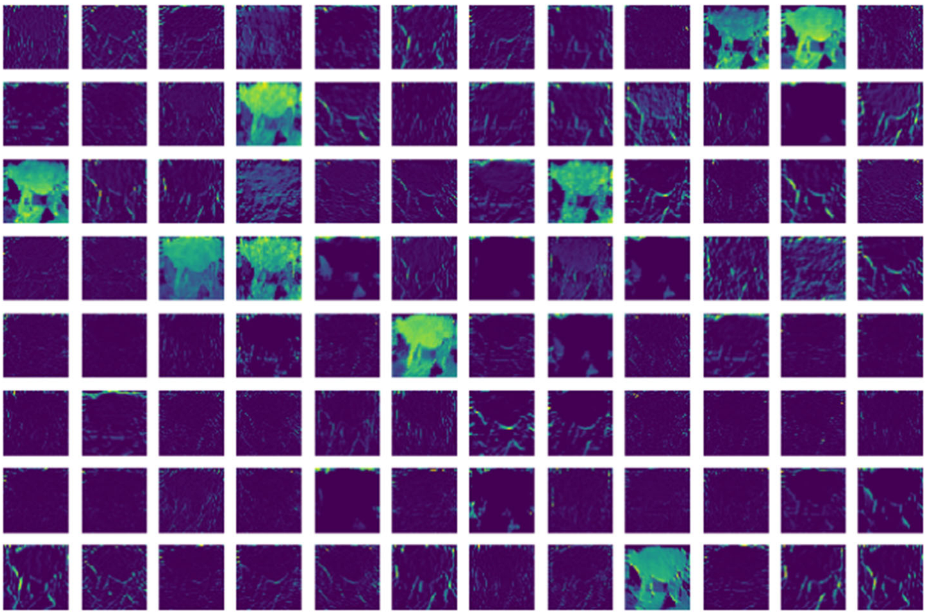

Fig. 5 Feature maps output by the first convolution layer

Fig. 6 Overlaying images of one input image and its corresponding feature maps

this experiment. We segmented the side-view images of the 93 cows. All of the areas of head, trunk and legs in those images were extracted. The identification performance of these three areas and the cow object are shown in Table 3.

As seen from Table 3, beyond the trunk area, the head and legs also contributed to the ability to identify the cows. The experiment also showed that the identification using the cow object was better than using partial areas. In this experiment, we realized the potential of combining each of the individual regional features for the individual identification. We will study cow identification based on the fusion strategy of deep part features in future work.

4.4 Misidentification analysis

Among the 418 validation images, 14 images were misidentified. Three typical examples of misidentification are shown in Fig. 8: cow No. 0 is identified as No. 1, cow No. 23 is identified as No. 3, and cow No. 84 is identified as No. 55.

By analyzing the examples of misidentification, it can be concluded that misidentification mainly occurs on three types of cow images: images of a cow whose main color is black, images of cows with small individual differences and blurry images. For the images of the cow whose main color is black, there is less individual information, and it is difficult to extract effective features. Therefore, it is easy for misidentification to occur. Markers could be added

Fig. 7 Part segmentation result

Table 3 Identification performance of three areas and cow object

Method	Accuracy (%)			
	head	legs	trunk	cow object
CNNs	80.51	86.67	97.44	97.78

to increase the individual feature information, to ensure that there is a certain degree of difference between similar black cows. Alternatively, deep edge features could be integrated to identify these cows. For the images of cows with small individual differences, fine-grained classification methods can be introduced in the identification process. For cow images with high ambiguity, such misidentification could be solved by selecting a camera with a high frame rate or using feature fusion (extracting features such as edge gray gradient and fusing them with the deep features).

5 Conclusion

In this work, we proposed an automatic algorithm for identifying cows using side-view images. The method used the YOLO object detection model to extract cow objects in the side-view images and used a fine-tuned AlexNet model to identify individual cows. We created an image dataset of 105 different cows, and the experiments were performed on this dataset. The proposed method obtained an accuracy of 96.65%, which has advantages compared to the traditional identification method and the deep learning method focused on a partial area. The experimental results showed that the method can be used to develop an automatic identification system of cows at farms, and an identification method based on deep learning has good application prospects and can provide accurate decision-making for precision livestock farming of dairy cows.

Fig. 8 Examples of misidentification (the top row is the validation sample, and the lower rows are examples of the identification results)

Acknowledgements This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFD0700204-02, in part by the Dong Nong Scholar Program of Northeast Agricultural University under Grant 17XG20, in part by the Natural Science Foundation of Heilongjiang Province of China under Grant QC2018074, in part by the Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs under Grant 2018AIOT-02, and in part by the China Agriculture Research System under Grant CARS-36.

References

1. Chen X, Li Q, Song Y, Jin X, Zhao Q (2012) Supervised geodesic propagation for semantic label transfer. In: European Conference on Computer Vision, 2012. Springer, 553–565
2. Girshick R (2015) Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, 2015. 1440–1448
3. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014. pp 580–587
4. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. pp 770–778
5. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: CVPR, 2017. vol 2. p 3
6. Jin X, Wu L, Li X, Chen S, Peng S, Chi J, Ge S, Song C, Zhao G (2018) Predicting aesthetic score distribution through cumulative jensen-shannon divergence. In: AAAI Conference on Artificial Intelligence (AAAI), New Orleans, Lousiana, USA
7. Johnston A, Edwards D, Hofmann E, Wrench P, Sharples F, Hiller R, Welte W, Diederichs K (1996) 1418001. Welfare implications of identification of cattle by ear tags. Vet Record 138(25):612–614
8. Kaixuan Z, He D (2015) Recognition of individual dairy cattle based on convolutional neural networks. Trans Chin Soc Agric Eng (Trans CSAE) 31(5):181–187
9. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, 2012. pp 1097–1105
10. Kühl HS, Burghardt T (2013) Animal biometrics: quantifying and detecting phenotypic appearance. Trends Ecol Evol 28(7):432–441
11. Kumar S, Tiwari S, Singh SK (2016) Face recognition of cattle: can it be done? Proc Ntnl Acad Sci, India Sect A: Phys Sci 86(2):137–148
12. Kumar S, Pandey A, Satwik KSR, Kumar S, Singh SK, Singh AK, Mohan A (2018) Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement 116:1–17
13. Li Q, Chen X, Song Y, Zhang Y, Jin X, Zhao Q (2014) Geodesic propagation for semantic labeling. IEEE Trans Image Process 23(11):4812–4825
14. Lin T-Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL (2014) Microsoft coco: Common objects in context. In: European conference on computer vision, 2014. Springer, pp 740–755
15. Lu H, Li B, Zhu J, Li Y, Li Y, Xu X, He L, Li X, Li J, Serikawa S (2017) Wound intensity correction and segmentation with convolutional neural networks. Concur Comput: Pract Exp 29(6):e3927
16. Lu H, Li Y, Uemura T, Kim H, Serikawa S (2018) Low illumination underwater light field images reconstruction using deep convolutional neural networks. Futur Gener Comput Syst 82:142–148
17. Lu H, Li Y, Mu S, Wang D, Kim H, Serikawa S (2018) Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. IEEE Internet Things J 5(4):2315–2322
18. Lu H, Li Y, Chen M, Kim H, Serikawa S (2018) Brain intelligence: go beyond artificial intelligence. Mob Netwo Appl 23(2):368–375
19. Redmon J, Farhadi A (2017) YOLO9000: better, faster, stronger. arXiv preprint
20. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. Proc IEEE Conf Comput Vis Pattern Recognit 2016:779–788
21. Ren S, He K, Girshick R, Sun J (2015) Faster r-cnn: towards real-time object detection with region proposal networks. Adv Neural Inf Proces Syst 2015:91–99
22. Serikawa S, Lu H (2014) Underwater image dehazing using joint trilateral filter. Comput Electr Eng 40(1):41–50
23. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556

24. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vis 57(2):137–154
25. Voulodimos AS, Patrikakis CZ, Sideridis AB, Ntafis VA, Xylouri EM (2010) A complete farm management system based on animal identification using RFID technology. Comput Electron Agric 70(2):380–388
26. Zhang L, Yang Y, Gao Y, Yu Y, Wang C, Li X (2014) A probabilistic associative model for segmenting weakly supervised images. IEEE Trans Image Process 23(9):4150–4159
27. Zhang Y, Gravina R, Lu H, Villari M, Fortino G (2018) PEA: parallel electrocardiogram-based authentication for smart healthcare systems. J Netw Comput Appl 117:10–16
28. Zin TT, Phyo CN, Tin P, Hama H, Kobayashi I (2018) Image Technology based Cow Identification System Using Deep Learning. In: Proceedings of the International MultiConference of Engineers and Computer Scientists

Weizheng Shen received the M.S. and Ph.D. degrees in agricultural electrification and automation from Northeast Agricultural University, Harbin, China, in 2003 and 2007, respectively. He is now a Professor in the School of Electrical Engineering and Information, Northeast Agricultural University, and a visiting scholar in computing and information science, Kansas State University. His research interests include computer vision, agricultural information technology and near infrared spectroscopy analysis.

Hengqi Hu received the B.Sc. degree in management from Heilongjiang Institute of Technology, Harbin, China, in 2017. He is currently pursuing the M.Sc. degree in computer science with the School of Electrical Engineering and Information, Northeast Agricultural University. His current research interests include computer vision and precision livestock farming, etc.

Baisheng Dai received his Ph.D. from School of Computer Science and Technology, Harbin Institute of Technology in 2017. He is currently a lecturer of School of Electrical Engineering and Information in Northeast Agricultural University. His current research interests include computer vision, precision livestock farming and intelligent agriculture, etc.

Xiaoli Wei graduated from the Northeast Agriculture University in 2005 with her master's degree, and received the Ph.D. degrees in Mechanical Design and Theory from Northeast Forest University in 2014. She is now an associate Professor in the School of Electrical Engineering and Information, Northeast Agricultural University. Her current research and work focus on Agricultural information technology, Wireless sensor network, and the related algorithm in artificial intelligence.

Jian Sun received her Bachelor's degree in industrial engineering and automation from Shenyang Polytechnic College (currently known as Shenyang Ligong University), Shenyang, China, in 1985. She is now an Associate Professor in the School of Electrical Engineering and Information, Northeast Agricultural University. Her research interests include computer vision, agricultural information technology and near infrared spectroscopy analysis.

Li Jiang received the M.Sc. degree from Northeastern University, Shenyang, China, in 2009. She is currently pursuing the Ph.D. degree in agricultural electrification and automation with the School of Electrical Engineering and Information, Northeast Agricultural University. She is currently a lecturer of School of Electrical Engineering and Information in Heilongjiang Bayi Agricultural University. Her current research interests include image processing and intelligent agriculture, etc.

Yukun Sun received the M.Sc. degree from Jilin Agricultural University, Changchun, China, in 2017. He is currently pursuing the Ph.D. degree with the School of Animal Science and Technology, Northeast Agricultural University. His current research interest is application of image information technology in dairy cow production.