Recognition method of dairy cow feeding behavior based on convolutional neural network

Zhefu Chen, Xiaodong Cheng, Xi Wang and Mingshu Han

School of Electronic Information Engineering, Inner Mongolia University, Hohhot 010000, China

Email: 1084655259@qq.com, cxd0808@imu.edu.cn, 122547880@qq.com, 657910531@qq.com

Abstract. With the continuous development of the Internet of things and artificial intelligence, smart animal husbandry has gained popularity in recent years. It can efficiently solve the labor cost and capital cost, and reduce the loss of pasture and herders. This paper proposes a method based on convolutional neural network to identify the feeding behavior of dairy cows, by detecting the feeding behavior of dairy cows, judging their health status, estimating feed intake, and estimating milk production according to a proportional coefficient. First, this article need go to the ranch to collect activity data and manually calibrate the cows; second, use the sliding window method to upgrade the one-dimensional activity data; finally, feed the two-dimensional data into the CNN, continuously adjust the parameters and lengthen data width, the recognition accuracy rate can reach 89.5%, realizing the recognition of the feeding behavior of dairy cows. The results show that the CNN-based dairy cow feeding behavior recognition method designed and implemented in this paper can realize efficient automatic recognition and has high accuracy.

1. Introduction

As far as the Inner Mongolia Autonomous Region is concerned, there are about 140 million cattle and sheep, and the total number of cattle is about 10 million. The demand for cattle and sheep is large, and the price has risen year by year. In the past five years, small families in pastoral areas have begun to merge and gradually form large-scale pastures. Therefore, it is impossible to observe by relying solely on manpower, and the labor cost is relatively high. At the same time, the Internet of Things and artificial intelligence technologies are developing rapidly, and policies are also vigorously supporting the development of smart animal husbandry, with a large investment. Therefore, the traditional industry (livestock) is improved by artificial intelligence methods, and the digitalization of the industry has become a trend.

Feed intake is the main factor influencing the growth and development of ruminant livestock and lactation. In the past ten years, my country's dairy industry has developed at an alarming rate. One of the main factors affecting the lactation of dairy cows is the intake of dairy cows. However, most of the current feeding detection methods target group dairy cows. However, accurate feeding data cannot be obtained for individual dairy cow feeding detection, and the nutritional status of individual dairy cows cannot be understood [1]. Accurately detecting the feeding behavior of individual dairy cows is the most important part of realizing the digital breeding of dairy cows. Therefore, the use of artificial intelligence methods to study the feeding behavior of individual dairy cows has become the focus of the intelligent dairy cow industry.
In addition, according to the individual food intake of each dairy cow, the health status of the dairy cows can be detected. If abnormalities occur, the dairy cows can be warned in time to remind herders to reduce the economic loss caused by the disease or death of the dairy cows. To estimate the feed reserves based on the intake of dairy cows and cost, avoiding the occurrence of insufficient feed or waste of feed. At the same time, the milk production of individual cows can be estimated based on the feed intake, which reduces the amount of human labor and saves human resources for large-scale farms. Realizing accurate and efficient breeding industry actively responds to the government’s policy on smart animal husbandry.

2. Proposed method

2.1 Neural network concept construction
In the late 1950s, F•Rosenblatt designed and produced the emerging concept of "perceptron", which is actually a neural network with a large number of layers. This is the first time in the world that artificial neural networks have been truly transformed from theoretical research to Practice process [2]. The "perceptron" itself is an algorithmic mathematical model that imitates the behavioral characteristics of animal neural networks and can process information in parallel and distributed. The algorithm used by the neural network is actually vector multiplication, plus the extreme value approximation of various functions. Inspired by its environment, input some data (text, graphics, etc.) to the network one after another, adjust the weights of each layer of the network according to actual requirements and different algorithms, and wait for the weights of each layer of the network to converge to a certain value. The learning process is over, and then we can use the generated neural network to solve classification and regression problems on real data.

2.2 Transfer learning
In real life, many problems, such as classification and regression problems, have insufficient data due to various reasons. At this time, transfer learning is the primary solution we consider, and its purpose is to solve the problem of insufficient training data in deep learning [3]. In transfer learning, the training model can be adjusted so that the data does not need to be trained from scratch, or the trained model can be borrowed from the source data set. The essence of transfer learning is feature reuse, which is manifested as parameter reuse; the characteristic is that tasks are similar; the types of transfer learning are: 1. Fully connected layer training under small data (at least the last fully connected layer of the network needs to be changed, and the model Perform fine-tuning, so that when the source data is small and the training network is not satisfied, avoid training from scratch), 2. Fine Tuning (selective training of convolutional layers other than the fully connected layer), 3. as Large-scale parameter retraining of model parameter pre-training (this step is performed when the amount of data is large).

This article is affected by the epidemic, and the data only has the one month feeding data of 6 cows before the epidemic. The amount of data involved in training is not large. Therefore, migration learning is a good method. Fully connected layer training under small data avoids training. In the process, the problem of neural network overfitting due to lack of source training data has been proven to be a good way to apply computer vision on small-scale data.

2.3 Convolutional Neural Network
Convolutional neural network is actually a feedforward neural network, one of the deep learning algorithms that contains convolution calculations and can contain deep structures. As shown in Fig.1: Its core has three main parts: 1. Convolutional layer 2. Introduction of nonlinear activation function 3. Pooling layer [3]. The above three parts are also called the "sandwich" structure of the convolutional neural network. The neural network in this article mainly includes three convolutional layers, two batches of normalization, a pooling layer, and a fully connected layer (acting as a classifier). The input data is the one-month feeding data of six black and white Holstein cows. Cov represents the convolutional layer, BatchNorm represents the batch standardization, and MaxPool
represents the maximum pooling. Finally, the fully connected layer is added and the Relu nonlinear activation function is used to play the role of return device, learning is complete.

The core part of CNN is the convolutional layer, and each convolutional layer contains many convolution kernels. In this paper, the one-dimensional data of the cow’s feeding is upgraded to two-dimensional data and then the convolution operation is performed, which is equivalent to generating the image of the feeding data. To perform convolution operations on its pixels (that is, the dot product between the convolution kernel and the pixels), as long as the width of the convolution kernel matches the data [4-6].

Define filter:
Enter the size: $H_1 \times W_1 \times D_1$ (1)

Four hyper parameters: Filter number K; Filter size F; Step size S; Zero padding size P.
Output volume: $H_2 \times W_2 \times D_2$ (2)

$$H_2 = \frac{H_1 - F + 2P}{S} + 1$$ (3)

$$W_2 = \frac{W_1 - F + 2P}{S} + 1$$ (4)

$$D_2 = K$$ (5)

The purpose of introducing the activation function is to introduce nonlinearity into the model. If there is no activation function, no matter how many layers the neural network has, it is a linear mapping. In this paper, instead of using the traditional tanh and sigmoid activation functions, the Relu function is selected as the activation function. The relu function acts as a return selector. Its calculation is simple and can speed up the model; due to the need for calculation bias in the back propagation process Derivative, most of the derivative points of the tanh function are zero, and only the input meets the value of that small area, the derivative is not 0; and the maximum value of the derivative of the sigmoid function is 0.25. If the sigmoid function is used, the back propagation of each layer will be Make the gradient at least a quarter of the original. When there are more layers, the gradient will disappear and the model cannot converge.

Define Relu function: $f(x) = \max(0, x)$ (6)

Neuron output: $\max(0, W^T + b)$ (7)

Pooling layer This article uses only the commonly used maximum pooling to eliminate other redundant data that affects the results, and retain the characteristic values of its own data rules (according to the invariance of local characteristics), reduce the number of parameters, and reduce computing time, improve the operation speed.
The reason for choosing batch normalization is that traditional normalization is to normalize the entire data set, while BN is to standardize small batch data (a batch). Because traditional normalization is to process all data, a certain batch or a certain part of the data may not meet the value we want, so BN makes the normalization more thorough, can promote the convergence of the loss function, and the model is more stable. In addition, for CNN, BN can well combat the problem of gradient disappearance, which can make CNN have hundreds of layers of networks without underfitting or gradient disappearance. After the epidemic is completely over, you can go to the cattle farm to collect more data. Then you need to deepen the network, and the effect of BN will be even more significant.

Traditional standardization: 
\[ x' = \frac{x - E[x]}{\sqrt{\text{Var}[x]} + \epsilon} \]  
(8)

Where \( E[x] \) represents the average value of neurons \( x \) in each batch of training data, and \( \text{Var} \) represents the variance.

BN: 
\[ y = \gamma x' + \beta \]  
(9)

\[ \gamma = \sqrt{\text{Var}[x]} \]  
(10)

\[ \beta = E[x] \]

\[ m = \sum_{i=1}^{m} x_i \] //mini-batch mean

\[ \sigma_B = \sum_{i=1}^{m} (x_i - \mu_B)^2 \] //mini-batch variance

\[ \bar{x}_i = \frac{x_i - \mu}{\sqrt{\sigma_B^2 + \epsilon}} \] //normalize

\[ y_i \leftarrow \gamma \bar{x}_i + \beta \equiv B_{N_{F_{\beta(x)}}} \] //scale and shift

In the above formula, \( m \) refers to mini-batch size.

2.4 CNN approximate equivalent RNN

Because the feeding behavior of dairy cows is actually a very regular time-series behavior, and the data is simple one-dimensional data, simple threshold condition judgments or recurrent neural networks (RNNs) are often used for such time-series one-dimensional data. But this article uses the CNN method. Although this method has advantages in processing images, it has a stronger ability in data fitting. This article deals with time data, so it is converted into a certain length of time series and fed to the network. From the actual network structure, it has been approximately equivalent to an RNN [7]. The reason why it is approximately equivalent is that the RNN processes data It is in a one-dimensional space, but this article deals with two-dimensional data, which can actually learn better features. Now companies will also use CNN for training text data (typical time series). The reason for using CNN is because CNN is convenient for parallel operations and has a short training time.
3. Hardware part

3.1 Composition of activity collector
When cows eat, their jaws move up and down, causing vibrations in the temporal fossa. There is obvious agitation in the temporal fossa position when the cow is feeding, and the number of agitations in the temporal fossa position is related to the number of chewing times of the cow. Basically, it is chewed once and the temporal fossa agitation; chewing 5 to 10 times will swallow once. Stop chewing when swallowing, and the vibration of the temporal fossa will stop [1,9].

Aiming at the feeding characteristics of dairy cows, the feeding activity collector is worn on the side of the cow’s neck near the jaw bone in the form of a collar, and the activity data of the dairy cow is collected 8 times per second. Activity collector mainly includes microcontroller (MCU), bubble activity sensor, ZigBee wireless transceiver module and power supply [14]. The microcontroller is an ARM-based ultra-low power 32-bit processor. The bubble activity sensor is a new type of ultra-low power activity sensor developed based on the principle of liquid sloshing. It consists of a transparent cavity, an infrared emitting diode and an infrared receiving diode. The transparent cavity contains a certain amount of silicone oil and air (called bubbles). When the sensor shakes, the movement of the air bubble will cause the change of the medium between the infrared emitting diode and the receiving diode, and the induced current of the infrared receiving diode will also change, so the movement trend is recorded as an electrical signal. The ZigBee wireless transceiver module works in the 2.4GHz frequency band and supports the secondary development of communication protocols [8].

![Fig.3 Bubble sensor](image)

3.2 Principle of base station
The main function of the base station is to receive the data sent by the collector, and then upload it to the cloud platform through the 4G-LTE transparent transmission module, or to read the data locally through the computer serial port. The base station mainly includes ZigBee wireless transceiver module, TTL to 232 level conversion circuit and 4G-LTE transparent transmission module. The base station also uses the ZigBee wireless transceiver module with a communication radius of about 80m, which can cover the test environment. There is a serial receiving software on the PC, which is used to receive and display the data sent by the base station, and perform necessary data cleaning, data filtering and adjustment on the data[10-13].

4. Data processing and experimental results

4.1 Data preprocessing
When collecting activity data of dairy cows, it is necessary to calibrate the real-time feeding situation of dairy cows as a comparison of subsequent model recognition results. At present, the more commonly used method is to arrange a camera on the scene to shoot, watch and analyze the video to get the feeding state of the cow. However, this method requires too much hardware, and the scope is limited, and sometimes it may not be possible to observe an individual accurately and carefully. In this paper, considering the field environment, we adopt the method of manually calibrating the collected
data through manual observation on the spot. The calibration status is the feeding and non-feeding status, the non-feeding status is calibrated as 0, and the feeding status is calibrated as 1. There is probably a range for cows’ feeding, which can be covered by numerical values. Set the upper and lower thresholds and filter out unreasonable numerical values. Finally, visualize these data to facilitate comparison with the subsequent on-site manual calibration data, and it is easy to find the law and cycle of feeding by cows, as shown in the Fig.4:

Because the base station uploads the data to the PC, the PC needs to design software to preprocess the received cow feeding data. This article uses the pascal language and the development environment for delphi7 to develop a simple software: serial communication controller. The data is set to be automatically saved every five minutes, and a txt format file is generated. Because the received data is a hexadecimal value, so the data needs to be converted from hexadecimal to decimal. The content of each column of the data is: year, month, day, hour, minute, Seconds, activity value, and the cow’s feeding state judged after identification, as shown in the Fig.5.

4.2 Data Upgrading
As mentioned above, because the cow feeding data collected by the collector is a one-dimensional value, and the convolutional neural network convolves a two-dimensional value, so this article uses the view operation in the Pytorch framework (reshape is not selected in the code, Instead, use view to reshape the dimensions of the data. Reshape will create a new tensor to allocate memory, and view will not reallocate memory. First, convert the data into a 2d tensor (for example, if the length of a one-dimensional data sequence is 36, change it 6*6), and finally fed into the neural network for training, as shown in the Fig.6.
4.3 Data distribution and comparison of experimental results

As shown in the Fig.7: The parameter w_width is an example of w_width = 11. This is the way to generate training data. Use this sliding window to scroll all data samples to create a data set, and then divide it into training and validation sets at a ratio of 3:1. The results come from the validation set. The data in the validation set is randomly selected and does not participate in training.

The results of this article are expressed in the form of accuracy: the number of samples to be paired divided by the number of all samples

\[
\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN}
\]

Using different methods, the same data and equipment, the specific results are as follows:

| Method          | Data dimension | Data width | Calculating speed | Maximum accuracy |
|-----------------|----------------|------------|-------------------|------------------|
| Threshold judgment | One-dimensional | 144        | 30 seconds        | 47.8%            |
| Neural Network   | Two-dimensional | 144        | 10 minutes        | 59.9%            |
| RNN              | One-dimensional | 144        | 50 minutes        | 72.4%            |
| CNN              | Two-dimensional | 144        | 20 minutes        | 89.5%            |

Obviously, when the data is the same, CNN has the highest accuracy.
When the deep learning methods are consistent, the wider the data width, the higher the accuracy rate, but the higher the hardware requirements of the corresponding calculation, the results of this article are run on the GPU. When the amount of collected data is sufficient, the data width can be enlarged, and the parameters can be adjusted continuously to improve the accuracy of recognition.

5. Conclusion
In deep learning, neural networks and their typical algorithms can effectively learn and train data to achieve ideal functions. For this reason, this paper proposes a method based on convolutional neural network to recognize the feeding behavior of dairy cows. Its innovation lies in the data itself, only one-dimensional activity value can reflect the characteristics of animal activity. Experimental results show that this method can more accurately identify the feeding behavior of dairy cows.

Acknowledgements
This work is supported by Inner Mongolia Natural Science Foundation of China (No. 2019 MS06011), and Hohhot City Applied Technology Research and Development Project (No. 2019-HT01).

References
[1] Zhou Yating. Research on Beef Cattle Feeding Behavior Recognition and Feed Intake Model[D]. 2018.
[2] Liu Ying, Zhang Qiannan, Dong Haitao, Cheng Mei, Liu Hongyan, Wang Fuping, Fan Jiulun. On-the-spot image classification method based on adaptive weight CNN feature fusion [C]. The 2017-2019 "Academic Golden Autumn" award-winning papers.: Shaanxi Institute of Communications, 2020: 7-11.
[3] Fu Yao, Wang Yajing, Cao Zhijun, Li Shengli. Research progress on factors affecting the feeding behavior of dairy cows [C]. Chinese Association for Science and Technology, Shaanxi Provincial People's Government. The 18th China Association for Science and Technology Annual Conference-14 Northwest Proceedings of the Symposium on the Development of Pasture and Animal Husbandry in Regional Agriculture and Animal Husbandry. China Association for Science and Technology, Shaanxi Provincial People's Government: Academic Department of Chinese Association for Science and Technology, 2016: 411-415.
[4] Juan C. Núñez, Raúl Cabido, Juan J. Pantrigo, Antonio S. Montemayor, JoséF. Vélez.Convolutional Neural Networks and Long Short-Term Memory for skeleton-based human activity and hand gesture recognition[J]. Pattern Recognition, 2018, 76.
[5] Xiaohui R, Gang L, Miao Z, et al. Dairy Cattle's Behavior Recognition Method Based on Support Vector Machine Classification Model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019.
[6] Gu Jiu-Xiang, Wang Zhen-Hua, Jason Kuen, et al. Recent advances in convolutional neural
networks. arXiv:1512. 07108v5, 2017.
[7] LU Lu, YANG Yuhong, JIANG Yuzhi, AI Haojun, TU Weiping. Shallow Convolutional Neural Networks for Acoustic Scene Classification[J]. Wuhan University Journal of Natural Sciences, 2018(02): 178-184.
[8] Wu H, Bai F, Zhong Z, et al. Research on Cow Rumination Monitoring Based on New Activity Sensor[C]// 2019 9th International Conference on Information Science and Technology (ICIST). 2019.
[9] Zhang Shuang. Research on the monitoring technology of individual dairy cow ruminating behavior [D]. Northeast Agricultural University, 2017.
[10] Tian Fuyang. Development of an automatic detector for dairy cattle feed intake [D]. Shandong Agricultural University.
[11] Wu Hongyu. Design and test of individual feed intake measuring instrument for dairy cows[D]. 2015.
[12] Liang Suzhen, Cai Zhenjiang, Wang Hui, et al. Dairy cow feed intake detection system based on virtual instrument[J]. Journal of Agricultural Mechanization Research, 2011, 33(002): 70-72.
[13] M. Steensels, A. Antler, C. Bahr, D. Berckmans, E. Maltz, and I. Ha-lachmi, “A decision-tree model to detect post-calving diseases based on rumination, activity, milk yield, bw and voluntary visits to the milking robot,” animal, vol. 10, no. 9, pp. 1493–1500, 2016.
[14] Rebecca, Handcock, Dave, et al. Monitoring Animal Behaviour and Environmental Interactions Using Wireless Sensor Networks, GPS Collars and Satellite Remote Sensing[J]. Sensors, 2009.