Clinical mastitis detection by on-line measurements of milk yield, electrical conductivity and deep Learn

Tian Fuyang¹, Wang Zhonghua², Yu Sufang³, Xiong Benhai⁴, Wang Shunxi⁵

¹College of Mechanical and Electronic Engineering, Shandong Provincial Engineering Laboratory of Agricultural Equipment Intelligence; Shandong Provincial Key Laboratory of Horticultural Machineries and Equipment, Shandong Agricultural University, Tai’an, China
²College of Animal Science and Technology, Shangdong Agriculture University, Tai’an, China
³College of Life Sciences, Shangdong Agriculture University, Tai’an, China
⁴Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China
⁵School of Engineering, China Agricultural University, Beijing, China
afytian@sdaau.edu.cn,  asdautfy@163.com
*sffy521@163.com

Abstract: Mastitis is the most common and costly disease in dairy cows since it can reduce milk yield, degrade milk quality, and increase healthcare costs. Detection of mastitis is an important part of udder-health management on dairy farms. Thus, the objective of this study is to develop a novel method for automatic on-line detection of clinical mastitis in an automatic milking system using the measurement of electrical parameters, data of milk production efficiency, and deep learning. The measurements were inputted into a neural network to calculate the mastitis detection index. The network was trained with 44 healthy and 6 clinical mastitic cows. 42 out of 44 healthy and 5 out of 6 mastitic cows were classified correctly after training. The trained neural network can predicted 164 out of 176 healthy quarters correctly in different evaluation data sets. These results were better than the results obtained with the model usually used on the farm.

1. Introduction
Increasing herd sizes in major dairying countries is driving a need for automation and technologies that help farmers with their daily management decisions¹⁻⁷. Cow mastitis refers to the inflammatory reaction caused by pathogenic microorganism infection in the breast, which is one of the most common cow diseases⁸. It is the disease with the highest incidence and the most serious harm among all cow diseases. Mastitis leads to the decrease of output and quality of dairy cows, the decrease of service life of dairy cows, and the huge economic loss. At present, about 220 million cows in the world lose as much as 35 billion US dollars every year due to mastitis, about 2 billion US dollars in the United States, and 3% of cows in Britain are eliminated due to mastitis. Depending on whether symptoms of clinical mastitis can be divided into two types of clinical mastitis and recessive mastitis⁹. Clinical mastitis has obvious clinical symptoms. Cow milk has visible changes, such as clumping and darkening of color, and the appearance
of lump and redness outside the breast. It is easy to be found and treated by taking immediate measures. However, due to the lack of any visible clinical symptoms, the diagnosis of invisible mastitis is relatively difficult. There is no obvious change in the appearance of the breast and milk, but there is a quantitative change in the chemical composition of milk. The specific manifestation is increased pH, increased electrical conductivity, and the number of somatic cells is more than 500,000/mL. Detection of hidden mastitis can control its diffusion and improve the treatment condition[9].

Somatic cell count (SCC) is the trait most commonly used in breeding programs for the analysis of mastitis resistance[10]. For the monitoring of SCC, DHI (dairy herd improvement) sampling analysis is mainly adopted, but the analysis cycle is long, the implementation process is cumbersome, the equipment is expensive, and the monitoring needs to be corrected frequently. Foss company of Denmark designed the FossMatic 5000 detector for somatic count, which is too expensive and difficult for general milk farms to afford, and requires professional personnel to operate. Compared with the former, PORTASCCC Digital Reader milk somatic cell detector in the United States is more convenient. This system is suitable for outdoor field detection, but the detection time is long, the storage requirements of test paper, activator and color card are relatively high, and the testing environment requires appropriate temperature and avoiding direct sunlight.

Many studies showed that EC of milk from cows affected by both clinically and subclinically mastitis is higher than EC of milk from healthy cows[11]. But previous studies have shown that the use of only EC in different detection algorithms was unable to achieve the ISO (2007) standard Se (>70%) and Sp (>99%) for CM detection[12].

The method of milk pH value is intuitive, easy to operate, and its cost is low, and it is not affected by the external environment temperature. Exploiting multisensor information could lead to sustainable improvements in the detection of mastitis.

Infrared thermograph technology (IRT) can sense the energy radiated by the thermal motion of the object itself, convert the Infrared signal into the processed electrical signal, to form the thermal Infrared image reflecting the surface temperature information of the object. By exploring the difference between eye surface temperature and breast surface temperature, IRT is used to detect mastitis of cows. Generally, the thermal imager is used to detect the eye surface temperature and breast skin surface temperature of dairy cows[13]. Due to the partial deformation of dairy cow breasts with the movement of dairy cows, it is impossible to achieve accurate positioning. Previous researches mainly focus on manual marking.

The purpose of this article is to establish the cow milk production, milk production efficiency, average electrical conductivity, high electrical conductivity, the milk (5 minutes), the highest average flow velocity of average flow velocity, pH, pH high temperature, milk, milk early days, whether suffering from mastitis, cow number 12 parameters and the cow estrus deep learning model of probability and use the model to validate the actual detection of milk.

2. Materials and methods
After infection with mastitis, chloride and sodium ions in milk increase significantly, which leads to the increase of milk conductivity and pH value. At the same time, the range of the milk conductivity of cows infected with mastitis was larger than the corresponding value without infection. So we through many experiments and based on analyzing the test results, find out the cow infected and uninfected mastitis milk electrical conductivity and pH value of the critical value and the change range of upper and lower critical point, and the critical value stored in the permanent memory of single-chip computer system as a future mastitis diagnosed benchmark, when the measurement of conductivity and pH value exceeds the specified benchmark, thought that the cow mastitis or suspected infection; On the contrary, it was considered not infected.

2.1. Data Source
From January to March 2019, data from 60 cows on a cattle farm in Tai’an, Shandong province, among them 50 holstein lactation cows were randomly selected and 10 mastitis disease cows. Data of 43,200 groups were collected from 4 lactation areas and 200 sampling sites for each cow. Milk data samples
were collected twice a day before entering the milking shed in the morning and evening. During the test period, 47 lactation cows were being healthy and 3 cows suffered mastitis disease.

2.2. Measurements
This detection sensor is mainly composed of the electrical conductivity measurement system, pH measurement system, temperature measurement system, clock system, communication system, signal detection, and control system (pH electrode, conductivity electrode, temperature sensors and corresponding modulation circuit), storage system, keyboard display system, dairy cattle identification system, expert system, and CPU, etc.

Because the conductivity and pH value varies with the temperature, a temperature measuring system is designed to compensate for the conductivity and pH value of milk at different temperatures.

Since the conductivity and pH value of milk of the same cow change with different fetal times, age, milking time, milking times, and lactation period, the critical points of conductivity, pH value, and range of conductivity for mastitis diagnosis should also be timely changed.

2.3. Data processing
The original dataset had nine measurements (features), the cow milk production, milk production efficiency, average electrical conductivity, high electrical conductivity, the milk (5 minutes), the highest average flow velocity of average flow velocity, pH, pH high temperature, milk, milk early days, whether suffering from mastitis, cow number.

As the dataset used for prediction of sub-clinical mastitis was collected from only a single farm, to provide a reliable base for generalization of the findings of this study, the dataset was transformed by Z-Standardization (for each feature, subtracting the mean and dividing by the standard deviation).

2.4. Deep Learn Algorithms
The classification algorithm examined in this study includes random forest (RF), support vector machine (SVM), k nearest neighbor (kNN), and adaptive boost (AdaBoost).

Random Forest[14] is an integrated learning method, which is composed of multiple decision trees trained on the training set. When applied to test data sets, the prediction of each tree model in a random forest is combined into a whole classification decision, for example, by majority voting or by applying weights. Because of this, the random forest model corrects the overfitting (if any) caused by the training set and provides robust classification performance.

Support Vector Machine (SVM) [15] is a nonprobabilistic classifier, which maps the input (i.e. classification feature) to a high-dimensional feature space, each dimension represents a classification feature. The support vector machine attempts to divide the high-dimensional feature space into two subspaces. The new now displayed data samples will be calculated later based on this partition to determine their class membership. This method is not probabilistic because the features in the unobserved data samples completely determine its position in the feature space of the support vector machine model.

KNN[16] is an instance-based non-parametric classification algorithm. The algorithm uses k nearest samples in training data as input to determine the membership degree of the class. Next, you assign the most common class membership between neighbors to the objects you want to classify. Generally, the distance between the objects of interest and their neighbors is weighted, so the neighbors who are closer to each other contribute more to the majority vote of class members.

Adaptive Boosting[17] is an integrated learning method, which determines the final output of the integrated classifier by a majority of votes. Boosting is a step-by-step process. One of the models is to train at each step. In each step, the weights assigned to training samples are modified to increase the weights of previously misclassified training samples and decrease the weights of correctly classified samples in the next step. AdaBoost supports a variety of machine learning algorithms as basic learners, but decision tree has proved to be a reliable and easy-to-use choice in the past.
2.5. Performance of the Classification

The performance of the classification algorithms was evaluated using the metrics of accuracy, precision, recall (also known as sensitivity), F-score and specificity, which can be computed as

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

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

\[
F - \text{Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}
\]

where \( TP \) (true positives) is the number of instances where a behaviour was correctly classified as the behaviour that was observed (ground truth). \( FN \) (false negatives) is the number of instances where a particular behaviour was observed (ground truth) but misclassified by the algorithm as some other behaviour. \( FP \) (false positives) is the number of instances where the algorithm falsely classified a behaviour that was not observed. \( TN \) (true negative) is the number of instances where behaviour was correctly classified as not being observed.

3. Results and discussion

Table 1 shows the data of healthy lactating cows numbered 16109 without mastitis. As can be seen from the data in the two tables, milk conductivity, pH value and milk-producing efficiency of cows have significant changes after mastitis. Due to the diversity of experimental data features, data with different features have different dimensions. In order to eliminate the differences in different data dimensions, experimental data need to be normalized. The dataset was transformed by Z - Standardization (for each feature, subtracting the mean and dividing by the standard deviation). Figure 1 shows the data curves of Z - Standardization processing various characteristics of the data after the offset.

| CO | Conductivity | Milk yield | pH  | Temperature | Milking time | Milking efficiency | Milk production duration | Milk production efficiency |
|----|--------------|------------|-----|-------------|--------------|-------------------|-------------------------|--------------------------|
| 161 16 | 9.43 | 6.8 | 6.04 | 37.93 | 4 | 1.275 | 12.031 | 0.424 |
| 161 16 | 10.2 | 5.1 | 6.18 | 38.28 | 5.4 | 1.148 | 12.127 | 0.511 |
| 161 16 | 9.99 | 6.2 | 6.22 | 37.82 | 5 | 1.02 | 11.844 | 0.431 |
| 161 16 | 9.95 | 5.1 | 6.29 | 37.71 | 5.2 | 1.192 | 12.036 | 0.515 |
| 161 16 | 10.01 | 6.2 | 6.1 | 38.07 | 4.5 | 0.867 | 11.802 | 0.33 |
| 161 16 | 9.91 | 3.9 | 6.04 | 38.26 | 6.6 | 1.152 | 12.323 | 0.617 |
| 161 16 | 10.25 | 7.6 | 6.09 | 37.66 | 5 | 0.72 | 11.684 | 0.308 |
| 161 16 | 9.95 | 3.6 | 6.22 | 38.03 | 4.7 | 1.064 | 12.397 | 0.403 |
| 161 16 | 8.95 | 5 | 6.31 | 37.51 | 4.6 | 0.717 | 12.011 | 0.275 |
Table 2 the evaluation of performance of the model

| Model               | Accuracy | Precision | Recall | F-score |
|---------------------|----------|-----------|--------|---------|
| Random Forest       | 92%      | 92.1%     | 76.8%  | 83%     |
| Support Vector Machine | 93.3%  | 92.6%     | 81.2%  | 85.1%   |
| k Nearest Neighbour | 94.6%    | 94.8%     | 87%    | 90.1%   |
| AdaBoost            | 92.3%    | 91.8%     | 80.7%  | 84.1%   |

Table 2 shows the evaluation of performance of the model based on the confusion matrix, the accuracy rate, precision rate, recall rate and precision rate of detection of cow mastitis based on kNN neural network in the existing samples were 94.6%, 94.8%, 87%, 90.1%, respectively. In order to illustrate the effect of lvq-based cow mastitis diagnosis scheme, this paper also compares the diagnosis methods based on random forest, support vector machine and BP neural network algorithm, and the results are shown in figure 5.
4. Conclusions
In this article, this study shows that applying deep learning to dairy products is also an interesting method to detect mastitis. The kNN deep learn model was used to classify dairy cow clinical mastitis, and the traditional support vector machine (SVM), Random Forest and adaboost model were compared. It can be seen from the test results that the kNN neural network based auxiliary diagnosis method for dairy cow clinical mastitis has a high accuracy rate. kNN neural network is expected to be an effective and practical auxiliary diagnosis tool for cow mastitis, reducing subjective misdiagnosis by computer and improving the diagnostic accuracy.

Acknowledgements
The authors acknowledged the financial support from the Chinese National Key Research & Development Plan (NO.2017YFD0701605-2, 2016YFD0700205), Shandong Province Key Technologies Research & Development Plan (NO.2017GNC10123) and Shandong Provincial Key Research and Development Plan (NO.2016CYJS05A02).

References
[1] Z. G. Yan, Z. H. Wang, Q. Zhang, S. J. Yue, B. Yin, Y. L. Jiang, K. R. Shi. Identification of whole-genome significant single nucleotide polymorphisms in candidate genes associated with body conformation traits in Chinese Holstein cattle. Animal Genetics. 2019. DOI:10.1111/age.12865
[2] S. J. Yue, Y. Q. Zhao#, X. R. Gu, B. Yin, Y. L. Jiang, Z. H. Wang*, K. R. Shi. A genome-wide association study suggests new candidate genes for milk production traits in Chinese Holstein cattle. Animal Genetics 2017, 48: 677-681 (doi: 10.1111/age.12593)
[3] Zhao XJ, Li ZP, Wang JH, Xing XM, Wang ZY, Wang L*, Wang ZH. Effects of chelated Zn/Cu/Mn on redox status, immune responses and hoof health in lactating Holstein cows. Journal of Veterinary Science, 2015, 16(4): 439-446.
[4] Jiang, SZ, Yang ZB, Yang WR, Li Z, Zhang CY, Liu XM, and Wan FC. Diets of differentially processed wheat alter ruminal fermentation parameters and microbial populations in beef cattle. Journal of Animal Science, 2015, 93: 5378-5385.
[5] Wang H., Yang W., Wang Y., Yang Z., Cui Y. The study on the effects of Chinese herbal mixtures on growth, activity of post-ruminal digestive enzymes and serum antioxidant status of beef cattle. Agricultural Sciences in China, 2011, 10(3): 448-455.
[6] Zhiyong Hu, Zhengyan Yin, Xueyan Lin, Zhengui Yan and Zhonghua Wang. Effects of feeding fatty acid calcium and the interaction of forage quality on production performance and biochemical indexes in early lactation cow. Journal of Animal Physiology and Animal Nutrition,2015,99(5), 899-904.

Fig.1 Dataset of the No. 16109 dairy cows after Z - Standardization
[7] Fugui Jiang, Xueyan Lin, Zhengui Yan, Zhiyong Hu, Yun Wang and Zhonghua Wang. Effects of forage source and particle size on feed sorting, milk production and nutrient digestibility in lactating dairy cows. Journal of Animal Physiology and Animal Nutrition, 2018, 102(6):1472-1481.

[8] Bradley, A. 2002. Bovine mastitis: An evolving disease. Vet. J.164:116–128.

[9] Cai Yixin, Ma Li, Liu Gang. Design and experiment of rapid detection system of cow subclinical mastitis based on portable computer vision technology[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017,33(Supp.1): 63－69. (in Chinese with English abstract)

[10] E. Norberg, Electrical conductivity of milk as a phenotypic and genetic indicator of bovine mastitis: A review, Livestock Production Science 96 (2005) 129-139.

[11] Cheng-Chang Lien, Ye-Nu Wan, Ching-Hua Ting. Online detection of dairy cow subclinical mastitis using electrical conductivity indices of milk[J]. Engineering in Agriculture, Environment and Food 2016, (9) 201-207

[12] M.Khatun, P.C.Thomson, K.L.Kerrisk, N.A.Lyons, C.E.F.Clark, J.Molfino, S.C.García. Development of a new clinical mastitis detection method for automatic milking systems[J]. Journal of Dairy Science, 2018, 101, (10): 9385-9395

[13] ZHANG Xudong, KANG Xi, MA Li1, LIU Gang. Automatic Detection Method of Dairy Cow Mastitis Based on Thermal Infrared Image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(7): 248 -255. (in Chinese)

[14] Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32.

[15] James, G.; Witten, D.; Hastie, T.; Tibshirani, R. An Introduction to Statistical Learning; Springer: Heidelberg, Germany, 2013.

[16] Altman, N.S. An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. 1992, 46, 175–185.

[17] Kégl, B. The return of AdaBoost.MH: Multi-class Hamming trees. arXiv 2013; arXiv:1312.6086.