In-line detection of clinical mastitis by identifying clots in milk using images and a neural network approach

G. Van Steenkiste[1,*], I. Van Den Brulle[1], S. Piepers[1] and S. De Vliegher[1]
[1]*Department of Internal Medicine, Reproduction and Population Medicine, Faculty of Veterinary Medicine, Ghent University, Belgium*
* Corresponding author: Glenn Van Steenkiste, glenn.vansteenkiste@ugent.be

Abstract

Automated milking systems (AMS) already incorporate a variety of milk monitoring and sensing equipment, but the sensitivity, specificity and positive predictive value of clinical mastitis (CM) detection remain low. A typical symptom of CM is the presence of clot(s) in the milk during pre-milking. The objective of this study is the development and evaluation of a neural network (NN) that can detect these clot(s) on pictures of the filters of the milking system after the pre-milking phase. The data for this study was generated by adding debris and/or clot(s) from used milk filters of AMS to milk and passing this milk through a blue circular milk filter mounted in a PVC tube. A camera was mounted in the PVC pipe to take a photo after each pass of milk. In total 696 pictures were taken with clot(s), and 586 pictures without. These were randomly divided into a 60/20/20 training, validation, and testing datasets, respectively, for the training and validation of the NN. A convolutional NN with residual connections was trained and the hyperparameters were optimized based on the validation dataset using a genetic algorithm. The integrated gradients were calculated to explain the interpretation of the NN. The accuracy, specificity, and sensitivity of the NN on the testing dataset were 100%. The integrated gradients showed that the NN identified the clot(s). Further validation and integration on farm in AMS are necessary, but the proposed method is very promising for the inline detection of CM on AMS farms.

Keywords: deep learning, mastitis, automated milking systems, image recognition

Introduction

From an economical view, mastitis is one of the most important disease in dairy cows (Halasa et al., 2007). Due to the effects on animal health and subsequent reductions in milk production and the need to discard abnormal milk or milk from diseased cows (European Union Directive EC/853/2004 and US Food and Drug Administration Grade A pasteurized milk ordinance), the cost of each clinical mastitis (CM) is approximately 485$. Early detection of CM can reduce both the economic impact and the long term impact on cow health and welfare (Heikkilä et al., 2012).

Many dairy farms are transitioning to automated milking systems (AMS) with around 38.000 units installed in 2017 (Sandgren and Emanuelson, 2017). When using AMS there are fewer opportunities for the farmer to detect CM in individual animals. The current AMS incorporate a variety of milk monitoring and sensing equipment, but the sensitivity and specificity of CM detection remain relatively low with most systems having a sensitivity between 47-90% and a specificity between 56-99% (Khatun et al., 2018; Penry, 2018). As a reference, the international standards organization (ISO) describes a standard target of 90% sensitivity and 99% specificity for the detection of abnormal milk (ISO/FDIS "Final Draft International Standard" 20966, Annex C "Automatic Milking Installations – Requirements and Testing"). Most of these sensor systems try to detect mastitis by measuring and analyzing indirect parameters such as (but not limited to) electrical conductivity, somatic cell count, milk flow rate, milk color, milk yield per hour or per quarter, and cow activity (de Mol and Ouweltjes, 2001; Fogsgaard et al., 2015; Khatun et al., 2018; Naqvi et al., 2022).

A typical symptom of CM is the presence of clot(s) in the milk during pre-milking, which has been proposed as the gold standard for the detection of CM (Mein and Rasmussen, 2008). Therefore we propose to use an in-line camera to detect these clot(s) in the filter after the pre-milking phase. A similar sensor has been proposed in the past, but was limited in its capabilities of detecting the clot(s) on the filter and was rather developed to score the quality of the milk (Wiethoff and Suhr, 2007). The objective of this study is the development and evaluation of a neural network (NN) that can detect these clot(s) on pictures of the filters of the milking system after the pre- and/or milking phase.

Materials and methods

Experimental data

The data for this study was generated by adding debris (straw, hay, manure, bedding material, mud, teat sealer, calcium, and flies) and/or clot(s) from used milk filters of AMS to milk and passing this milk through a circular milk filter (Universal Hygia Favorit filters, Universal dairy equipment) mounted in a PVC tube. The filters were painted blue for better visualization of the clot(s). An iPhone 6s was mounted in the PVC pipe to take a photo with the flashlight after each pass of milk. In total 696 pictures were taken from filters with clot(s), and 586 pictures from filters without clot(s).

Image analysis

For the training dataset the images without clot(s) were randomly resampled to obtain an even number of images with and without clot(s) for balancing the NN weights. During the training of the NN the images were randomly rotated, flipped, rescaled, zoomed, and sheared. In order to avoid that the NN learned features from outside of the filter, e.g. milk spatters on the PVC pipe, the PVC pipe was removed from the picture using OpenCV v4.1. The image was then rescaled to 500*500 pixels as NN input (Figure 1).

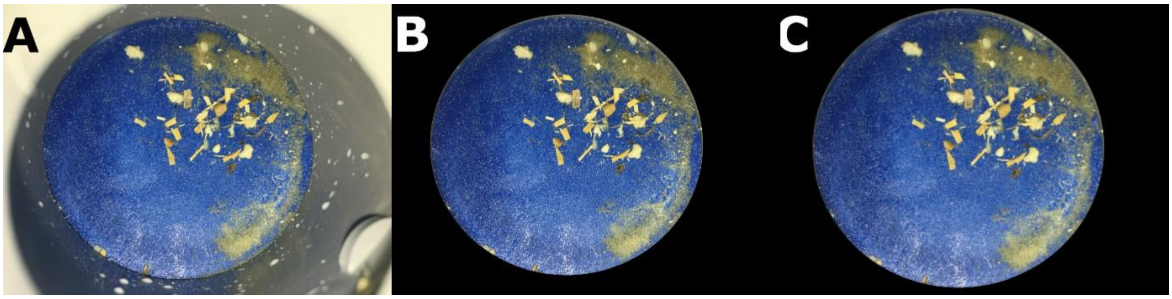

Figure 1: Image pre-processing steps on filter image after the passing of milk with clot(s). Panel A: original image taken wit a resolution of 4032*3034 pixels. Panel B: image after applying a black mask on the region of the PVC pipe. Panel C: resulting image used for the neural network after rescaling to 500*500 pixels.

A convolutional NN with residual blocks was designed for the image classification (Figure 2). The first parallel path inside a residual block consisted of 2 2D convolutional layers, followed by a batch normalization layer, a rectified linear linear unit activation (ReLU) block and a dropout layer (van Steenkiste et al., 2020). Zero-padding was used after convolution in order to keep the temporal order of the input signal. The second path of the convolutional block subsamples the input using a maximal pooling operation with the same subsample factor as the combined convolutional layers of the first part of the path (He et al., 2015). In total 8 residual blocks were used sequentially in the network. After the residual blocks the results are flattened to a single dimension by concatenating and fed into a dense (fully connected) layer after batch normalization, ReLU and dropout. An additional dense layer with ReLU and dropout was added and connected to a final dense layer with soft-max activation in order to produce a distribution over the 2 output classes (clot(s) or no clot(s)).

A genetic algorithm was used to optimize the hyperparameters of the NN based upon the validation dataset. A non-dominated sorting genetic algorithm II (NSGA-II) was used with a population size of 20 and 10 optimization generations with the accuracy as fitness value (Deb et al., 2002). Training of each child NN of the algorithm was ended after 50 epochs or when the training stopped improving for 3 epochs. Optimization was used for the following hyperparameters: number of filters, width of convolution and subsampling for each individual convolutional layer, number of neurons for each fully connected layer, L2 regularization and dropout used for training. For training the SoftMax cross entropy was used to calculate the loss with the Adam optimizer with the default parameters with a learning rate of 0.0001 for updating the weights of the network (Kingma and Ba, 2014). The network was built in Keras with TensorFlow backend (Chollet, 2015). Batch size was set to 16. Afterwards, the integrated gradients of the NN were calculated compared to a completely black baseline image in order to obtain an insight into the input-output behavior of the neural network. The attribution of the input pixels to the output labels was projected as a mask over the input image using the OpenCV toolbox (Figure 3).

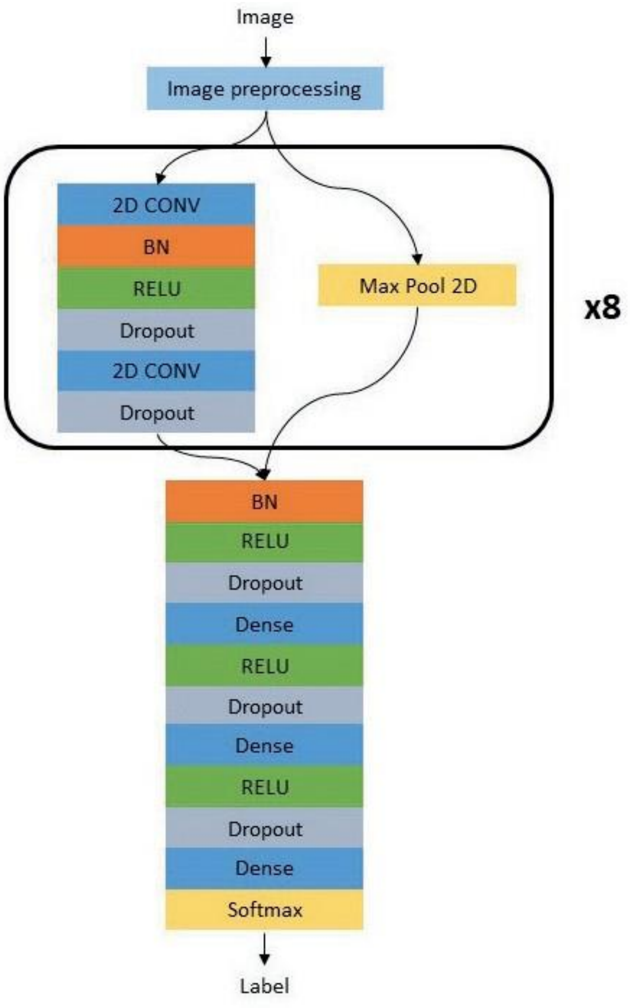

Figure 2: High level architecture of the neural network. CONV: convolutional block; BN: batch normalization; RELU: rectified linear activation unit; Max Pool 2D: Max pooling operation; Dense: fully connected layer; Softmax: Softmax activation block with the 2 different output classes.

Statistical analysis

The dataset was randomly divided into a 60% training subset of data, 20% validation, and 20% holdout subset which was not used for tuning the model, resulting in 1006 training, 335 validation, and 335 holdout images. The following metrics were calculated on the holdout dataset: accuracy, positive- and negative predictive value, specificity and sensitivity.

Results and discussion

The accuracy, positive- and negative predictive value, specificity, and sensitivity of the NN on the holdout dataset were 100%, which is a clear improvement in comparison with other milk based CM detection methods. Since the proposed method can reliably detect clot(s) in pre-milk this method can be a feasible approach to detect CM on AMS. Since clot(s) during pre-milking are considered the gold standard for detecting CM, the current approach has the potential to be a more reliable CM detection implementation for AMS in comparison with current CM detection sensors that try to detect CM with other parameters (Mein and Rasmussen, 2008; Penry, 2018).

One of the main benefits of the current approach is the extremely high accuracy. The main frustration of dairy farmers is the current high number of false alarms by the available CM detection sensors on AMS (Mollenhorst et al., 2012). Even if the practical implementation of the current sensor would not have an accuracy/precision of 100%, a NN can be adapted in order to maximize the accuracy by penalizing false positive results during the training process or by calculating the receiver operating characteristic curve and setting a manual threshold for the minimal required specificity (Fawcett, 2006). An additional benefit of using NN is their robustness for the presence of a variety of detriments (such as straw, manure, udder or tail hair, sawdust, sand, remainder of internal teat sealants) on the image without the need to retrain the algorithm for every possible detriment. This is in contrast with the previous proposed work for which the fuzzy logic algorithm had to be adapted for recognizing the different detriments (Wiethoff and Suhr, 2007). If environmental changes (e.g. change of filter type) would overwhelm the robustness of the NN, the weights of the NN could be updated on site with transfer learning to adapt to the new environment (van Steenkiste et al., 2020). If the farmer would receive multiple false positive results of the sensor, he could initiate an update of the algorithm based upon the incorrectly classified images without the needs to re-engineer the algorithm.

If a 3D representation of the filter would be created, e.g. by adding a second camera for 3D stereovision the NN could also be adapted to calculate the volume of the clot(s). Calculation of the animals clot(s) volume allows to monitor diseased animals with clinical mastitis (e.g. recovery and severity of the disease) over time and to estimate the severity of the disease. For example, if the clot(s) volume of a diseased animal is decreasing between consecutive milkings, the animal is recovering. Finally, if only a few clot(s) are present, increasing the number of milkings per day can be a sufficient treatment. This will contribute to a decrease in antibiotic use. On the other hand, treatment can be necessary when many clots in milk are detected.

The integrated gradients showed that the NN identified the region of the clot(s) as input feature (Figure 3). Neural networks are notorious for being a "black box" approach. It is difficult attributing the prediction of a NN to its input features and thus know why a NN told us which pixels of an image are responsible for picking a certain label (Sundararajan et al., 2017). The aim of explainable artificial intelligence (AI) is to understand the input-output behavior of NN. One of such explainable AI methods is the integrated gradients approach in which the attribution of each pixel is calculated by summing the gradients (partial derivative of each variable, while all others are held constant) of the network on different points at the path between a baseline image (e.g. a black image) to the actual input image (e.g. the image of the filter with clot(s)). If the PVC pipe had not been removed from the input images, the integrated gradients would have clearly shown that the NN had learned to recognize the different milk spatter patterns instead of recognizing the clot(s).

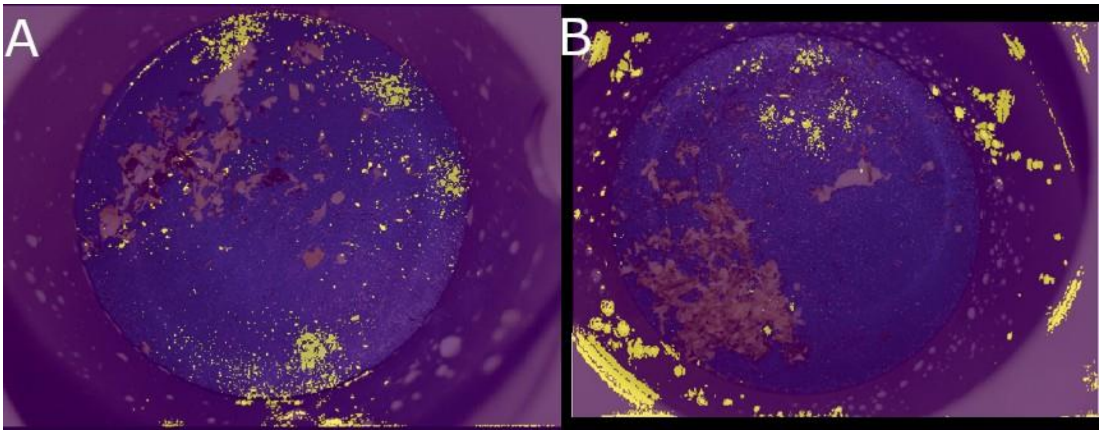

Figure 3: Visualization of the integrated gradients by an attribution mask over the original input image. Yellow indicates a high attribution to of the indicated pixels to the output label of the neural network (NN). Panel A shows the integrated gradients of the currently used NN, here the pixels (around) the clot(s) attribute the most to the output label. Panel B shows the integrated gradients of a NN which was trained on images were the PVC pipe was not removed from the image. This 'cheating' NN clearly used the milk spats on the PVC pipe to identify if this was an image with clot(s).

Conclusions

The current paper proposed an inline CM detection sensor for AMS. Further validation and integration on farm in AMS are necessary, but the proposed method is very promising for the accurate inline detection of CM on AMS farms.

References

Chollet, F. (2015) Keras. *https://keras.io*.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2), 182-197.

de Mol, R.M., and Ouweltjes, W. (2001) Detection model for mastitis in cows milked in an automatic milking system. *Preventive Veterinary Medicine* 49(1–2), 71-82.

Fawcett, T. (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27(8), 861-874.

Fogsgaard, K.K., Bennedsgaard, T.W., and Herskin, M. S. (2015) Behavioral changes in freestall-housed dairy cows with naturally occurring clinical mastitis. *Journal of Dairy Science* 98(3), 1730-1738.

Halasa, T., Huijps, K., Østerås, O., and Hogeveen, H. (2007) Economic effects of bovine mastitis and mastitis management: A review. *Veterinary Quarterly* 29(1), 18-31.

Heikkilä, A.M., Nousiainen, J.I., and Pyörälä, S. (2012) Costs of clinical mastitis with special reference to premature culling. *Journal of Dairy Science* 95(1), 139-150.

He, K., Zhang, X., Ren, S., and Sun, J. (2015) Deep Residual Learning for Image Recognition. *IEEE Conference On* 32(5), 428-429.

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

Kingma, D.P., and Ba, J. (2014) Adam: A Method for Stochastic Optimization. *ICLR 2015*, 1-15.

Mein, G.A., and Rasmussen, M.D. (2008) Performance evaluation of systems for automated monitoring of udder health: Would the real gold standard please stand up. *Mastitis Control: From Science to Practice*, 259-265.

Mollenhorst, H., Rijkaart, L.J., and Hogeveen, H. (2012) Mastitis alert preferences of farmers milking with automatic milking systems. *Journal of Dairy Science* 95(5), 2523-2530.

Naqvi, S.A., King, M.T.M., Matson, R.D., DeVries, T.J., Deardon, R., and Barkema, H.W. (2022) Mastitis detection with recurrent neural networks in farms using automated milking systems. *Computers and Electronics in Agriculture* 192, 106618.

Penry, J.F. (2018) Mastitis Control in Automatic Milking Systems. *Veterinary Clinics of North America - Food Animal Practice* 34(3), 439-456.

Sandgren, C.H., and Emanuelson, U. (2017) Is there an ideal automatic milking system cow and how is she different from an ideal parlor milked cow. In: *National Mastitis Council 56th Annual Meeting* St. Pete Beach, Florida, USA, 201761-68.

Sundararajan, M., Taly, A., and Yan, Q. (2017) Axiomatic Attribution for Deep Networks. In: *34th International Conference on Machine Learning* Sydney, Australia, 5109-5118.

van Steenkiste, G., van Loon, G., and Crevecoeur, G. (2020) Transfer Learning in ECG Classification from Human to Horse Using a Novel Parallel Neural Network Architecture. *Scientific Reports* 10(1), 186.

Wiethoff, M., and Suhr, O. (2007) Method and device for determining the quality of milk produced by machine milking. *https://patents.google.com/patent/US8261597B2/en.*