# Milky WAI: Unlocking the Secrets of Raw Cow Milk through Speckle Pattern and AI

Valentina Bello
*DIII, University of Pavia*
Pavia, Italy
0000-0003-3231-0604

Irene Bassi
*DIII, University of Pavia*
Pavia, Italy
0009-0004-7554-8377

Matteo Fiocchi
*DIII, University of Pavia*
Pavia, Italy
0009-0004-6941-0953

Cristina Nuzzi
*DIMI, University of Brescia*
Brescia, Italy
0000-0001-5530-6136

Paolo Bellassi
*DiSTAS, Università Cattolica del Sacro Cuore*, Cremona, Italy
0000-0001-9528-2239

Luca Cattaneo
*DIANA, Universita Cattolica del Sacro Cuore*, Piacenza, Italy
0000-0001-6027-7536

Simone Pasinetti
*DIMI, University of Brescia*
Brescia, Italy
0000-0002-5098-6395

***Abstract*—This paper introduces a novel method for raw cow milk classification combining speckle pattern (SP) imaging and AI-based processing of statistical parameters. Raw cow milk has been classified considering (i) 6 statistical features extracted from raw SP images, and (ii) 4 features extracted from the Gray-Level Co-occurrence Matrix (GLCM) computed on each SP image. We conducted 4 experimental campaigns, resulting in $24,000$ SP images retrieved from 20 milks produced by 20 cows. We then considered two different datasets: (i) a *Complete* dataset (made of all $24,000$ frames), and (ii) a *Reduced* dataset, built by excluding data from one acquisition campaign. The two datasets were then split into training and testing sub-datasets. We trained three Wide Neural Networks (WNN) using different strategies. WNN1 and WNN2 have been trained and tested using the *Complete* and the *Reduced* datasets respectively, while WNN3 has been trained on the *Reduced* tested on the *Complete* dataset. WNN3 only attained $72\%$ accuracy (revealing sensitivity to environmental conditions), while WNN1 and WNN2 achieved a test accuracy higher than $90\%$, demonstrating effectiveness without relying on computationally expensive models such as CNNs. This approach could be considered as a promising initial step for cow milk classification using SP imaging.**

***Index Terms*—Data analysis, machine learning, measurement science, milk analysis, speckle pattern.**

## I. Introduction

Milk is a highly nutritious food consumed all over the world. Also, milk (especially bovine milk) is utilized as the base for a variety of dairy products such as cheese, yogurt, cream, and regional culinary specialties. In 2023, India emerged as the largest global consumer of cow milk, with consumption exceeding $87,000$ million liters. Following closely was the European Union, consuming $23,700$ million liters of milk. While cow milk production worldwide has experienced a modest increase over the last five years, maintaining an annual output of approximately $549,000$ million liters, the biggest producer in 2023 was the European Union, with just under $105,000$ million liters, with the United States ranking second in production [1]. In Europe, the major cow milk producers are Germany, France, the Netherlands, Italy, and Poland, accounting for 22%, 16%, 10%, 9%, and 9% of the total European production respectively [2]. Milk composition, however, is not always the same: the quantity of fats, proteins, and lactose, as well as other nutrients, can change depending on cows' genotype, lactation stage, and dietary regime [3]. Environmental conditions change milk composition as well, as studied in [4] where the authors focused on the effect of higher temperatures.

As a consequence, it is important for farmers to determine raw milk composition to eventually adjust their feeding schedules (e.g. a lack of nutrients can result in poor quality of the milk). This is generally conducted through complex laboratory chemical analyses. Known chemical methods include the Geber method and the Rose-Gottlieb method for fat content detection, the Kjeldahl method for protein quantity detection, and a variety of enzymatic methods to detect the quantity of lactose [5]. Since these methods require specialized laboratories and personnel, involve pre-treatment of the sample, and are time- and money-consuming, they are not suitable for frequent on-site testing of milk samples. The research community has developed many alternative detection techniques over the years, including chromatography, spectroscopy, measurements using the fluid dielectric properties, and novel sensing devices leveraging optical, physical, or chemical properties of the fluid under testing [5]. One of the most adopted technologies is NIR and MIR spectroscopy [6], [7]. Furthermore, several commercial devices are available to test milk samples with good accuracy, the most known ones are Milkoscan (Foss produced by FOSS Analytical), DairySpec FT (Bentley Instruments), and Milk Analyzers (Ekomilk Horizon). They are based on IR spectroscopy coupled with Fourier transform (named FTIR in short), thus they are quite expensive and their size makes them not suitable for portable low-energy usage.

Among optical, non-invasive, low-cost techniques one of the most promising is speckle pattern (SP) imaging. This method relies on the interaction of coherent light with either a rough surface or an opaque fluid. In particular, in the case of milk, light is back-scattered by the lipid micelles and globular proteins which are in suspension in a liquid matrix of water,

This project was supported by the European Union program "Next Generation EU, M4C2 component, investment 1.1", issued as an open call for the national project "PRIN PNRR 2022", grant agreement N° P2022FX228 - MUSKETEER: Milk adUlteration detection using SpecKlE paTtern and machinE lEaRning.

glucose, and soluble proteins. An interference pattern is then generated and it can be acquired by a camera (in the form of images) and analyzed to extract valuable information. SP generated by a liquid is particularly challenging to analyze since it is time-variant because of the Brownian motion of suspended particles (dynamic SP). The SP is strongly sensitive to the properties of the medium, such as its composition, temperature, density [8], [9]. Hence, by analyzing the properties of SP images, it is possible to infer information about the fluid's characteristics.

Traditionally, SP imaging has been used to measure the roughness of surfaces, thermal strain, and elastic modulus of mechanical specimens. However, more recently it has been exploited to test several fluids, including milk, coupled with statistical feature extraction to categorize samples and differentiate among them, or also to detect their adulteration (both fraudulent and accidental) that correlates with food safety [10]. For instance, in [11], the authors demonstrated SP application to study the fermentation process of milk according to temperature variations by means of image-based features of interest extracted from the SPs. Statistical features were adopted in [12] instead, proving that they may be sufficient to differentiate between samples of rice milk diluted with different concentrations of water. Advanced techniques involve the adoption of machine learning and deep learning to categorize a wider variety of samples that may differ very little from each other. In [13], the authors adopted a pre-trained Convolutional Neural Network (CNN) to classify milk samples according to their fat content. Furthermore, particle suspensions prepared ad-hoc as powders [14] or microplastics [15] can also be analyzed by SP analysis coupled with CNNs. The main difference between the adoption of CNN-based and feature-based techniques is that the model employed is computationally expensive in the former case, with the advantage of requiring little to no pre-processing since the CNN takes as input the SP directly without further analysis. In contrast, feature-based methods require a pre-processing step to extract features and determine which ones are the most important to describe the fluid characteristics, but the model is lightweight and easy to deploy on cost-effective devices.

In this work, we propose SP imaging in combination with machine learning to automatically classify raw cow milk samples produced by different cows, which mainly differ from each other in their nutritional composition. SP images of each milk sample were acquired in 4 different campaigns to evaluate also how different testing conditions affect the samples' composition over time, thus making the model robust towards small changes in the sample structure.

## II. Materials and Methods

### A. Instrumental configuration

The instrumental configuration exploited to carry out experimental measurements on milk samples is reported in Fig. 1. A semiconductor laser diode (L658P040, Thorlabs, NJ, USA) is used as a coherent light source to irradiate the samples. It emits a maximum optical power of 40 mW at the wavelength of 658 nm. It is supplied through a current driver (LDC500, Thorlabs, NJ, USA) and connected to a temperature controller (PRO800, Thorlabs, NJ, USA) for thermal stabilization. An aspheric lens with an anti-reflection coating (C230260P-B, Thorlabs, NJ, USA) is used to focus the radiation onto the surface of a transparent plastic cuvette (of volume 4.5 mL, and size $10 \times 10 \times 50$ mm) containing the sample, at an angle of about 30°. A linear polarizer (LPVISE100-A, Thorlabs, NJ, USA) is positioned before the lens to select only the main polarization component. Images and videos of the SP generated by the scattering samples are acquired using a monochrome CMOS camera (Alvium 1800 U-158m, Allied Vision, Germany) located in front of the cuvette. The selected orientation of the camera prevents the Snell reflection from the cuvette wall from impinging on and saturating the CMOS sensor. The camera sensor has a size of $5.02 \times 3.75$ mm, a total number of pixels of $1456 \times 1088$, and a pixel size of $3.45 \times 3.45$ $\mu$m. A second linear polarizer is placed before the CMOS sensor. The camera is connected via USB to a laptop, allowing data acquisition through proprietary software (Vimba Viewer, Allied Vision, Germany).

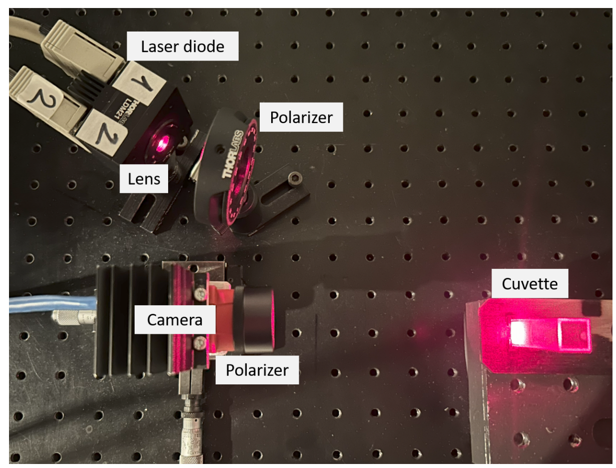


Fig. 1: Picture of the experimental setup.

### B. Data Acquisition

Raw milk samples were collected from 20 different cows (labeled with a numeric ID) by the *CERZOO - Centro di Ricerche per la Zootecnia e l'Ambiente Research Center* (Piacenza, Italy). For each sample, fats, proteins, and lactose quantities were measured using MilkoScan.

The raw samples were then moved to the laboratory for analysis. SP data were recorded on the same day of milking, and the morning after. The laser current was set to 85 mA (resulting in an optical power of about 23 mW reaching the sample). The camera exposure time and frame rate were set to 800 $\mu$s and 50 fps. A total of 4 Campaigns were conducted on each sample with different measuring conditions:

1) **Campaign 1:** samples temperature set to T = 22 °C, measurements conducted in the morning of the first day of testing.

2) **Campaign 2:** samples temperature set to T = 22 °C, measurements conducted in the afternoon of the first day of testing.
3) **Campaign 3:** samples temperature set to T = 10 °C, measurements conducted in the afternoon of the first day of testing.
4) **Campaign 4:** samples temperature set to T = 22 °C, measurements conducted in the morning of the second day of testing.

For each campaign, we prepared a fresh cuvette of the samples. We acquired 3 sets of 100 frames SP images for each sample analyzed in every campaign, for a total of 6 000 frames per campaign (100 frames $\times$ 3 sets $\times$ 20 samples). This results in a dataset of $24,000$ frames ($6,000$ frames per campaign $\times$ 4 campaigns). A sample SP obtained from one sample is shown in Fig. 2

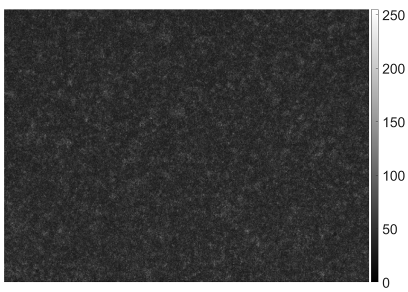


Fig. 2: Example of an SP image of a raw milk sample. Gray levels span from 0 (black) to 255 (white).

## III. Proposed Methodology

### A. Features extraction

Feature-based AI approaches are based on a pre-processing step to extract relevant information from the raw data, thus facilitating the consequent analysis. In the case of SP image analysis, this step is fundamental since analyzed SP images are monochromatic images from which the fluid's characteristics cannot be directly derived using image-based models such as CNNs. The analysis and machine learning model were both conducted using MATLAB 2023b.

From each SP image, we extract a number of features as described below. They were selected according to existing literature [12].

*1) Statistical features:* Typical statistical features that are extracted from SP images are (i) the image mean intensity in the range [0 - 255], (ii) the standard deviation of the image intensity in the range [0 - 255], (iii) the image kurtosis, (iv) the image skewness, (v-vi) the Full Width at Half Maximum (FWHM) of the autocovariance function of the SP image, which results in two values referring to the vertical and horizontal dimensions of the SP grains (FWHMx and FWHMy, respectively).

*2) Gray-Level Co-Occurrence Matrix features (GLCM):* The GLCM of an image represents the distribution of co-occurring pixel values at a given offset. It computes how often pairs of pixels with a specific intensity value and offset occur in the image [10]. GLCMs are often used to measure the texture of images and are a powerful tool in SP analysis. Given their definition, GLCMs are sparse matrices so a set of parameters can be extracted from them to be used as features for the machine learning models. In this work, we extracted the contrast, correlation, energy, and homogeneity parameters of the GLCM matrix.

TABLE I: Datasets composition in detail.

| Data | $\mathbf{Complete_{tr}}$ | $\mathbf{Complete_{te}}$ | $\mathbf{Reduced_{tr}}$ | $\mathbf{Reduced_{te}}$ |
|---|---|---|---|---|
| **Considered campaigns** | 4 | 4 | 3 | 3 |
| **Samples** | 20 | 20 | 20 | 20 |
| **Sets per sample** | 2 | 1 | 2 | 1 |
| **Frames per set** | 100 | 100 | 100 | 100 |
| **Total frames** | 16000 | 8000 | 12000 | 6000 |

### B. Datasets creation

Given the different sample temperatures that characterize Campaign 3 ($T_3 = 10$ °C with respect to $T_1$, $T_2$, and $T_4$ = 22 °C), we created two datasets: dataset *Complete* in which all the data is stored, and dataset *Reduced* in which data from Campaign 3 were removed. Both datasets were split into training and testing sub-datasets. To do so, instead of randomly selecting the samples following the usual $80\%-20\%$ proportion, we randomly selected for each campaign and each sample a single random set of $100$ speckle patterns. In this way, only two sets of $100$ frames each for every sample of every campaign were used to construct the training set, resulting in the data division shown in Table I, in which subscripts "tr" and "te" refer to training and testing sub-datasets respectively.

### C. Machine learning analysis

Among the plethora of models available for training in MATLAB "Statistics and Machine Learning" toolbox, we adopted the preset Wide Neural Network model. The network had only one fully connected layer at the end, a first layer size of $100$ neurons with ReLU activation function, and an iteration limit of $1,000$. We chose this model after a preliminary investigation of the performance of our dataset on the available models in the toolbox. Moreover, neural networks typically achieve high prediction accuracy compared to other models, especially when a lot of data is provided as input (both in terms of quantity of samples and quantity of features). The model input variables were the $6$ statistical features detailed in Section III-A1 and the $4$ GLCM features detailed in Section III-A2, while the model output was the cow IDs.

We implemented three different models combining available datasets:

1) **Experiment 1 (E1):** obtained by training the network on the $Complete_{tr}$ sub-dataset and testing it on the $Complete_{te}$ sub-dataset. This produced WNN1.
2) **Experiment 2 (E2):** obtained by training the network on the $Reduced_{tr}$ sub-dataset and testing it on the $Reduced_{te}$ sub-dataset. This produced WNN2.

3) **Experiment 3 (E3):** obtained by training the network on the $Reduced_{tr}$ sub-dataset and testing it on the $Complete_{te}$ sub-dataset. This produced WNN3.

E3 has been defined to analyze the models' robustness. Since the $Complete_{te}$ dataset includes data from Campaign 3 excluded from the $Reduced_{tr}$ dataset, the goal of E3 was to highlight the importance of including data from the same samples acquired under different environmental conditions, such as temperature and time passed from milk's collection (which may induce separation between its contents and eventually fermentation). For each experiment, model validation was conducted using a cross-fold of 5 layers.

## IV. Results and Discussion

The network performed well in both E1 and E2, achieving a test accuracy of $91.7\%$ for WNN1, and an accuracy of $91.6\%$ for WNN2. The test accuracy values between the two models are not surprising since they were evaluated on the same kind of data that they were trained to interpret. Referring to Fig. 3 and Fig. 4 (showing the confusion matrices obtained in E1 and E2), it is evident that WNN1 misclassifies more often the data with other classes (e.g. the number of yellow-colored cells in the confusion matrix is higher than in WNN2); however, the number of wrong predictions in each yellow cell is generally lower, the highest being $11.5\%$ in the case of samples of Cow 9 wrongly predicted as being of Cow 10. In contrast, WNN2 has higher error rates but in fewer classes (e.g. the number of yellow-colored cells is lower but their values are higher), the higher being $16.3\%$ in the case of samples of Cow 18 misclassified as Cow 19. This difference can be easily grasped by looking at the true positive rates (TPR) and false negative rates (FNR) on the right of the confusion matrixes, which summarize results by row. In the case of WNN1, TPR values below $90\%$ are achieved for samples of Cows 2, 4, 9, 10, 11, and 19 with FNR values ranging from $12.7\%$ to $19.2\%$. However, in WNN2 samples of Cow 2, 3, 4, 11, and 18 achieve TPR values below $85\%$, with a couple of them even below $80\%$, with FNR values ranging from $17.0\%$ to $23.7\%$. This result might not be acceptable depending on the particular application, as it indicates a higher chance of misclassification for these data compared to WNN1. A possible explanation for this difference is due to the dataset composition. In the case of E1, which included data from all Campaigns, the model confuses the samples with other ones more often since the numerosity of the dataset is higher. The data from Campaign 3 is quite different from the others due to sample temperature and this could be the reason why some samples are misclassified. In addition, the similar lactation stage and nutritional composition of some milk samples results in similar SPs, so the features are quite similar.

On the other hand, WNN3 achieves only a $72.0\%$ test accuracy. This result proves that adding data from campaigns conducted with different environmental conditions improves the robustness of the model, thus making it able to classify samples in more operating conditions. Referring to Fig. 5 (reporting the confusion matrix obtained in E3), it is worth noting that samples of Cow 12 were always predicted correctly, however when analyzing data of other samples the model often misclassified them as Cow 12. This highlights that Cow 12 data acquired in Campaign 3 may be similar to those of other samples acquired in the other Campaigns. This can be explained by considering that the molecules in colder fluids move slower, thus the SPs have different statistical characteristics compared to those acquired when the sample temperature was higher. The other two samples with higher error rates (that may be due to the same temperature effect) are Cow 4 misclassified as Cow 10 and Cow 7 misclassified as Cow 18. Therefore, the TPR values are always below $75.0\%$ (except for data of Cow 12 which has a TPR of $100\%$) and the FNR values range from $25.5\%$ to $37.2\%$.

## V. Conclusions

This paper presents a novel method to automatically classify raw cow milk samples. The approach is based on the acquisition of images of SP and subsequent application of artificial intelligence (AI) algorithms on 10 statistical features extracted from SP data (i.e. mean light intensity level and its standard deviation, kurtosis, skewness, FWHMx and FWHMy of the autocovariance function, GLCM contrast, GLCM correlation, GLCM energy, GLCM homogeneity). A total of 20 raw milk samples produced by different cows were analyzed, differing from each other by their nutritional content. We conducted 4 experimental campaigns to collect data from the samples at different environmental conditions (e.g. time passed from milk collection, sample temperature), obtaining a total of $24,000$ frames. We divided the data into two datasets named *Complete*, containing all the data collected, and *Reduced*, in which data from Campaign 3 was removed. The two datasets were split into training and testing sub-datasets by randomly removing one set of 100 frames from each sample in the dataset. The model chosen for the analysis was a Wide Neural Network coded in MATLAB. We trained 3 variations of the same model: WNN1, obtained by training and testing on data from the *Complete* dataset, WNN2, obtained by training and testing on data from the *Reduced* dataset, and, finally, WNN3, obtained by training on the *Reduced* dataset and testing on the *Complete* dataset.

The first two models achieved similar high accuracy values above $90.0\%$, demonstrating that the chosen features allow for the accurate classification of raw milk samples without employing complex models such as CNNs, which are computationally expensive and need specialized hardware to function properly. WNN3 achieved an accuracy of just $72.0\%$, highlighting that different environmental conditions can highly affect the characteristics of the fluid being tested, and therefore the features extracted from SP images. Hence, to improve the model robustness, it is highly recommended to include in the training dataset samples tested at different environmental conditions.

In conclusion, these results highlight that SP measurements combined with AI-based analyses are a cost-effective solution to grasp even subtle differences in fluid samples compared to

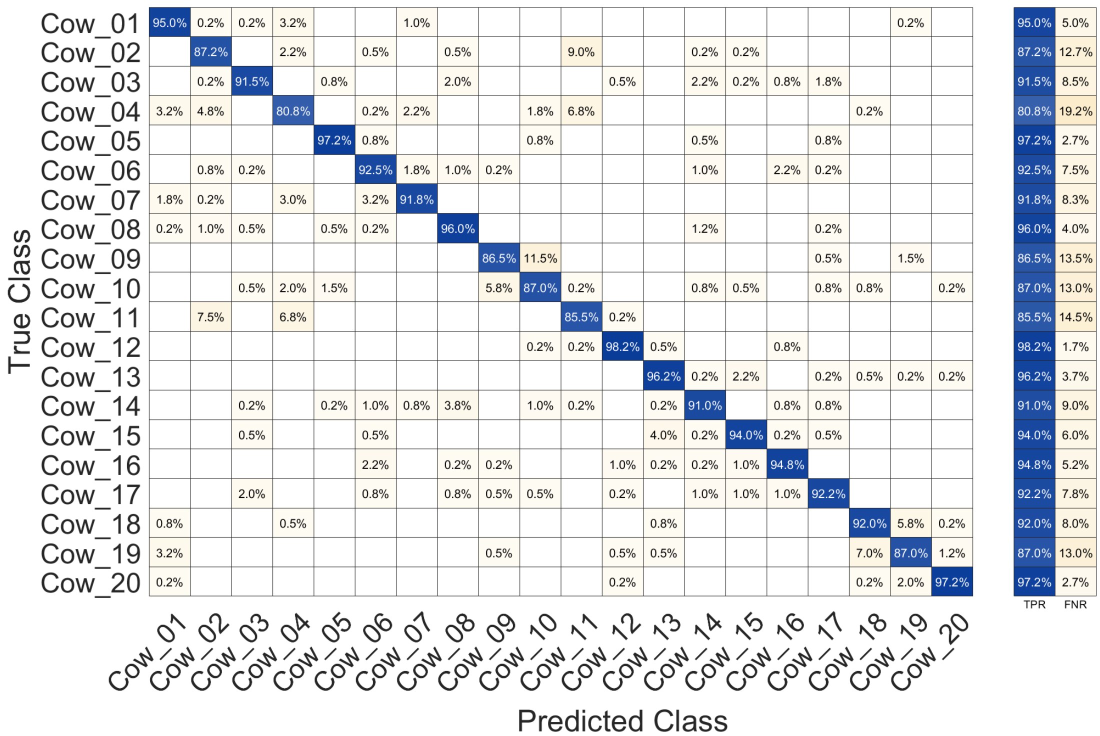


Fig. 3: Confusion matrix obtained in E1.

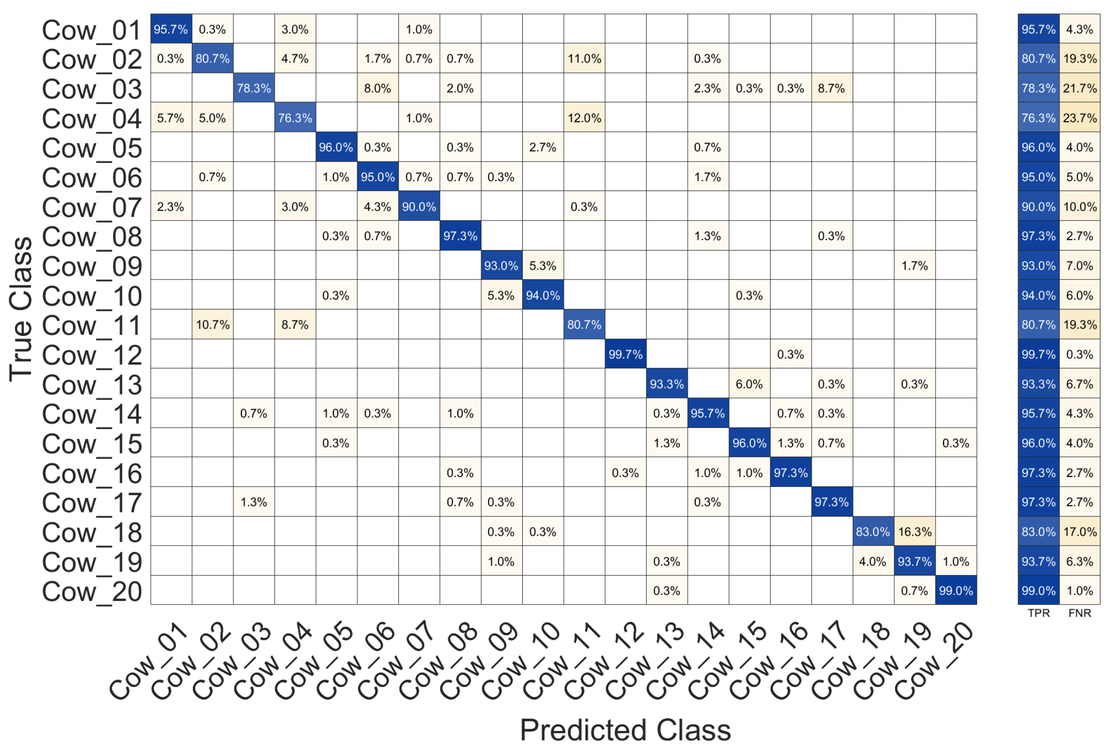


Fig. 4: Confusion matrix obtained in E2.

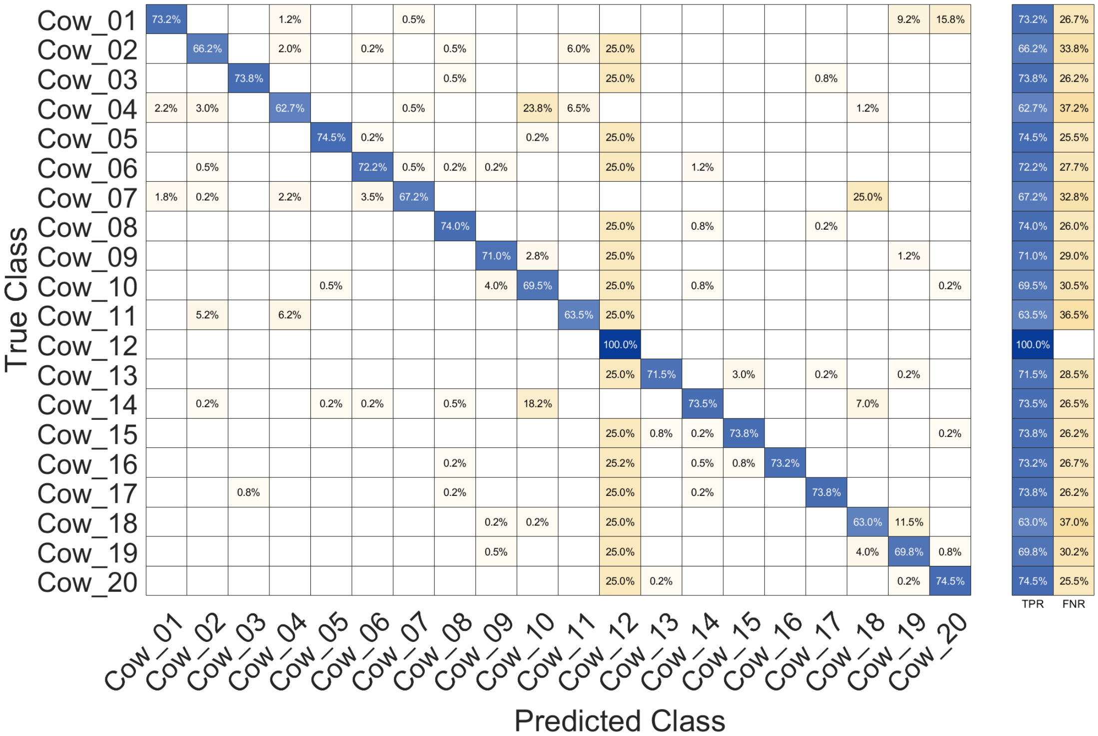


Fig. 5: Confusion matrix obtained in E3.

more expensive techniques such as spectroscopy and chemical analyses. The consequent analysis does not require computationally expensive hardware to be equally effective, in contrast with sophisticated approaches based on CNNs. Despite this, the proposed approach could be considered as an initial step for cow milk classification using speckle pattern images. In future advances, we aim to develop new approaches to improve the classification results and explore regression models to infer specific quantities of the sample's contents (e.g. the quantity of fats, proteins, and lactose in milk samples).

## Acknowledgments

The authors wish to thank Prof. Sabina Merlo for her precious guidance. The authors are grateful to Prof. Giuseppe Martini for his invaluable insights.

## References

[1] USDA Foreign Agricultural Service, "Annual consumption of cow milk worldwide 2023," https://www.statista.com/statistics/272003/global-annual-consumption-of-milk-by-region/, 2023, accessed: February 27, 2024.

[2] CLAL, "Eu-27: milk production and population," https://www.clal.it/en/?section=produzioni_popolazione, 2023, accessed: February 27, 2024.

[3] E. Vanbergue, L. Delaby, J. Peyraud, S. Colette, Y. Gallard, and C. Hurtaud, "Effects of breed, feeding system, and lactation stage on milk fat characteristics and spontaneous lipolysis in dairy cows," *Journal of Dairy Science*, vol. 100, no. 6, pp. 4623–4636, 2017.

[4] U. Bernabucci, S. Biffani, L. Buggiotti, A. Vitali, N. Lacetera, and A. Nardone, "The effects of heat stress in italian holstein dairy cattle," *Journal of Dairy Science*, vol. 97, no. 1, pp. 471–486, 2014.

[5] Z. Zhu and W. Guo, "Recent developments on rapid detection of main constituents in milk: a review," *Critical Reviews in Food Science and Nutrition*, vol. 61, pp. 312–324, 2021.

[6] R. Tsenkova, S. Atanassova, K. Itoh, Y. Ozaki, and K. Toyoda, "Near infrared spectroscopy for biomonitoring: cow milk composition measurement in a spectral region from 1,100 to 2,400 nanometers1," *Journal of Animal Science*, vol. 78, no. 3, pp. 515–522, 2000.

[7] H. Soyeurt, P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler, "Estimating fatty acid content in cow milk using mid-infrared spectrometry," *Journal of Dairy Science*, vol. 89, no. 9, pp. 3690–3695, 2006.

[8] J. W. Goodman, *Statistical Properties of Laser Speckle Patterns*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1975, pp. 9–75.

[9] S. C. Schneider, S. J. Rupitsch, and B. G. Zagar, "Signal processing for laser-speckle strain-measurement techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2681–2687, 2007.

[10] R. Pandiselvam, V. Mayookha, A. Kothakota, S. Ramesh, R. Thirumdas, and P. Juvvi, "Biospeckle laser technique – a novel non-destructive approach for food quality and safety detection," *Trends in Food Science Technology*, vol. 97, pp. 1–13, 2020.

[11] K. Chhanda and K. N. Anil, "Assessment of milk fermentation process at three temperatures using laser speckle technique," *Laser Physics*, vol. 32, no. 4, p. 045602, 2023.

[12] V. Bello, E. Bodo, and S. Merlo, "Speckle pattern acquisition and statistical processing for analysis of turbid liquids," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–4, 2023.

[13] K. Nyandey and D. Jakubczyk, "Combining transmission speckle photography and convolutional neural network for determination of fat content in cow milk: an exercise in classification of parameters of a complex suspension," *The European Physical Journal Plus*, vol. 139, pp. 123–137, 2024.

[14] J. Yan, M. Jin, Z. Xu, L. Chen, Z. Zhu, and H. Zhang, "Recognition of suspension liquid based on speckle patterns using deep learning," *IEEE Photonics Journal*, vol. 13, no. 1, pp. 1–7, 2021.

[15] D. Endo, T. Kono, Y. Koike, H. Kadono, J. Yamada, and U. M. Rajagopalan, "Application of laser speckles and deep learning in discriminating between the size and concentrations of supermicroplastics," *Optics Continuum*, vol. 1, no. 11, pp. 2259–2273, 2022.