

# Artificial intelligence to boost traceability systems for fraud prevention in the meat industry

Alessandro Biglia,<sup>1</sup> Paolo Barge,<sup>1</sup> Cristina Tortia,<sup>1</sup> Lorenzo Comba,<sup>1,2</sup> Davide Ricauda Aimonino,<sup>2</sup> Paolo Gay<sup>1</sup>

<sup>1</sup>Department of Agricultural, Forest and Food Sciences (DiSAFA), University of Turin, Grugliasco (TO);

<sup>2</sup>Institute of Electronics, Information Engineering and Telecommunications (IEIIT - CNR), Turin, Italy

## Abstract

Traceability was introduced about twenty years ago to face the worldwide spread of food safety crises. Traceability data flow associated with each lot of food products during any production and/or delivery phases can also be used to guarantee product authenticity. For this purpose, it is necessary to protect the data from cyber intrusions and, at the same time, to guarantee the integrity of the bond between the physical product and the data. Price grading related to quality perceivable or credence attributes attracts criminals to attempt item substitution fraud. Improved track and trace technologies supported by artificial intelligence

(AI) could highly enhance systems' capability to detect authenticity violations by product substitution.

This paper proposes an innovative method based on AI, to reinforce traceability systems in detecting possible counterfeiting by product substitution. It is an item-based mass balance method that analyses the congruity of the traceability data flows not by using explicit (even stochastic) rules but by exploiting the learning capabilities of a neural network. The system can then detect suspect information in a traceability data flow, alerting a possible profit-driven crime. The AI-based method was applied to a pork slaughtering and meat cutting chain case study.

Correspondence: Lorenzo Comba, Department of Agricultural, Forest and Food Sciences (DiSAFA), University of Turin, Largo Paolo Braccini 2, 10095 Grugliasco (TO), Italy.  
E-mail: lorenzo.comba@unito.it

Key words: Meat industry; food fraud; food counterfeiting; deep learning; mass balance.

Acknowledgements: the authors would like to acknowledge the Research Centre on Animal Production (Reggio Emilia, Italy) for providing data. Also, the authors would like to thank M.A.A. Franchetto for collaborating with this research work.

Funding: this work was supported by the project 'PININ - PiemuNt chènNa' (POR FESR 2014-2020 - Piattaforma Bioeconomia - Piedmont Region).

Contributions: Dr. Davide Ricauda Aimonino is co-last author of the paper.

Conflict of interest: the authors declare no potential conflict of interest.

Received for publication: 19 December 2021.

Accepted for publication: 5 August 2022.

©Copyright: the Author(s), 2022

Licensee PAGEPress, Italy

Journal of Agricultural Engineering 2022; LIII:1328

doi:10.4081/jae.2022.1328

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Publisher's note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

## Introduction

Nowadays, food counterfeits, fraud, and fakes are increasing, especially for high-end products such as wine, extra-virgin olive oil, ripened cheese, meat, *etc.*, resulting in reputation and economic losses. Therefore, it is recognised that a shift in focus from intervention to prevention should be made to reduce the size of the problem (Spink, 2019; Spink *et al.*, 2019).

Traceability tools can be exploited to prevent, deter and eliminate illegal, unreported and unregulated food production. Indeed, traceability procedures are fundamental in maintaining the food chain integrity, which concerns aspects of food production, such as the way food items have been sourced, procured, and distributed. In this case, food fraud is related to quality attributes that are not perceivable by the consumer (Robson, 2021). Also, regarding food authenticity, defined as the matching between the food product characteristics and the corresponding food product claims, a high level of efficiency in tracking and tracing food items is required both during the production processes as well as throughout the supply chain logistics (Dabbene *et al.*, 2014; Barge *et al.*, 2020). Furthermore, honest and accurate food labelling is a requirement that legal authorities must ensure. Therefore, robust and reliable analytical methodologies must be adopted to detect infringements (Kendall *et al.*, 2019; Wadood *et al.*, 2020; Esteki *et al.*, 2021).

The capability of traceability procedures in preventing fraud in a supply chain derives from the ability to trace the history, process, and location of an entity through recorded identification and the unique and inviolable coding and identification of the products to avoid possible infiltrations, swapping, or mixing of unauthorised products. In addition, the availability of data stored in different registries, web portals, and information exchange platforms allows the analysis of supply chain fluxes, thus preventing larger-scale food fraud (Ulberth, 2020). Fraud prevention and anti-counterfeiting can also be performed by overt (visible) and covert (difficult or impossible to see with naked eyes) product authentication

technologies, which, in any case, should be paired with methods for tracking and tracing products movements through the supply chain (Spink *et al.*, 2017; El Sheikha, 2021; Hellberg *et al.*, 2021).

The main actions to oppose counterfeits by tracking systems are the assurance of the link between physical items and digital twins and the integrity control of the physical flow of food masses. Therefore, several solutions to guarantee the identification code have been developed, such as non-removable labels (with holograms, RFID, optical codes, *etc.*), anti-tampering seals, *etc.* (Kemény and Ilie-Zudor, 2016; Barge *et al.*, 2017; Soon and Manning, 2019). In some cases, this can be very difficult to implement due to the nature of the product (Barge *et al.*, 2019) or the peculiarity of the operations conducted in the production stages (Comba *et al.*, 2011). This is the case for the example reported in this paper, which considers pork thighs. The pig's genetic type and weight at slaughtering determine the quality of the thighs and, consequently, of the ham. Indeed, the meat of some pig genetic types, characterised by a faster growth rate (and thus usually cheaper), may contain an excessive quantity of fat that is inadequate for high-quality cured ham production. The opportunity to substitute high-quality pork thighs with cheaper ones from other pig breeds could be very attractive to fraudsters. To detect such fraud, analytical control of the main commercial pig genetic types is possible by employing, for example, DNA-based methods which have been developed significantly, at lower cost and enhanced speed of execution (Wilkinson *et al.*, 2012; Galimberti *et al.*, 2013; Böhme *et al.*, 2019). The availability of analytical techniques is very powerful and, in many cases, is the only accepted legal proof in forensic debates. However, track and trace technologies coupled with artificial intelligence methods consent to analyse a wider spectrum of data flows through paths even characterised by physically distant nodes (international trade). In several supply chains, the production process along the whole path through primary production, breeding, logistics, and processing plants, is regulated by specific rules which protect quality and marks. This is the case, *e.g.* of protected designation of origin (PDO) and protected geographical indication (PGI), which in Italy cover more than 300 registered products (data updated in 2021).

Since January 2020, in certified Italian PDO ham value chains, checking the conformity to the voluntary certification schemes

relies on registering the regulations compliance of the animals in the farms and during transport (farm-to-farm, farm-to-slaughterhouses, and slaughterhouse-to-cutting plants) on a website (R.I.F.T., *Registro Italiano Filiera Tutelata*). Through this system controls, certain and reliable information from the birth of the pigs until the certification of the thighs, is shared. The pigs' birth month and weight data are registered by breeders, slaughterhouses, and cutting plants and cross-checked by the certification authorities and ham protection consortia.

The elaboration of the data arising from certified, shared, real-time transactions combined with mass balance-based algorithms can contribute to fraud prevention in several food supply chains. Recently, blockchain applications have been proposed to guarantee the inviolability of food traceability information (Xu *et al.*, 2020; Dey *et al.*, 2021; Niknejad *et al.*, 2021). Traceability models using product volume and mass balances are also employed to check the integrity of organic and sustainable food value chains (Mol *et al.*, 2015). In mass-balance models, the traded volume of certified sustainable produce is monitored throughout the entire value chain to ensure that the volume of the certified products downstream equals the volume of the certified resource base upstream of the same value chain. However, the mass-flow control may be vulnerable whenever a product (or a lot) is processed and split into many sub-products (or sub-lots), possibly generating some waste that is discarded. The same applies when bulk products are managed without a strict segregation policy (Comba *et al.*, 2013). Typically, since weighing is time-consuming, only the most valuable parts (or sub-lots) are singularly weighed. This makes mass balances between inputs and outputs unreliable, if not entirely inapplicable. Hence, it would be difficult to detect one or more parts being swapped in this context.

The hereby proposed artificial intelligence method was developed to verify the congruity of the mass balance of the raw material between the input and the output of the process by considering the true yields given by the production process, in the case of split mass flows, with only partial information (Figure 1). The data should be acquired and sent to the neural network whenever lots are split or joined (Dupuy *et al.*, 2005) or, more generally, whenever the granularity (*i.e.*, the size of a traceability unit) in the production/delivery flow is changed (Karslen *et al.*, 2011). Indeed,

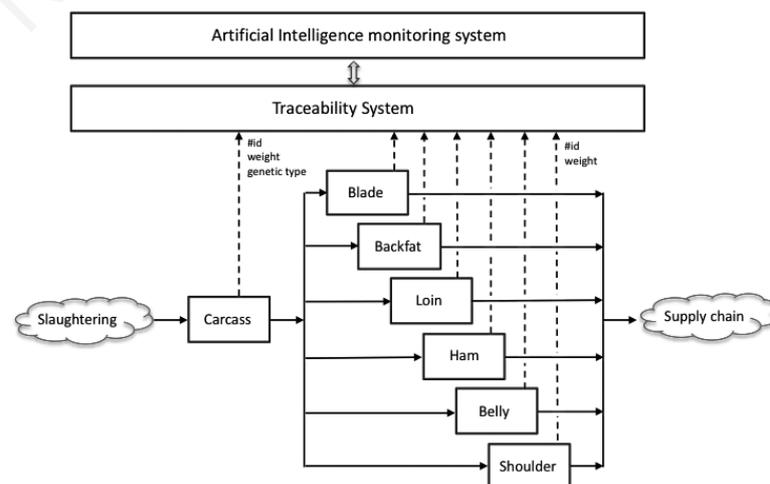


Figure 1. The integrated traceability and artificial intelligence system applied to the slaughtering and cutting process.

opportunities for fraud occur when the performance of a traceability system decreases, the granularity worsens and/or systematic information loss takes place, for instance, if the information about the process parameters is not properly linked to the product and systematically recorded (Dabbene *et al.*, 2014). The point where this loss occurs has been defined by Karlsen *et al.* (2011) as the critical traceability point (CTP). The identification and mapping of CTPs are typically performed by qualitative methods (direct observation, structured interviews, and document analysis) and lead to the definition of a CTP analysis plan. These are the points that should be monitored in the adopted approach by the NN-based method presented in this paper. Indeed, while this type of control is very simple in basic supply chain schemes (*i.e.*, few partners, well-established and linear links), the supervision of huge supply chains could be very difficult.

Moreover, to be competitive in the market, firms may outsource phases of the production process with the consequence that food system frameworks become more complicated as the number of links increases. In this context, only intelligent systems that can learn from large amounts of data involving many supply chain stakeholders can detect possible inconsistencies even in real-time. With the same principle, artificial intelligence techniques have detected potential abnormal activities such as financial fraud (Al-Hashedi and Magalingam, 2021) and administrative errors (Young *et al.*, 2021).

The originality of this paper is the adoption of a neural network coupled with a traceability system to verify the integrity of the joint product-information flow. Indeed, research is increasingly developing new tools able to exploit neural networks to improve food items' traceability. For example, Wang *et al.* (2017) connected a neural network to a traceability system to classify the quality of products employing a fuzzy algorithm, but they did not identify possible fraud. Vo *et al.* (2020) also used a neural network in conjunction with image processing techniques to enable an automatic grading solution in the southern rock lobster supply chain. A novel method based on deep learning to detect possible fraud in ginger or turmeric powder was proposed by Jahanbakhshi *et al.* (2021).

The paper is structured as follows: the proposed method and its application to the case of a pig slaughtering and cutting chain to detect possible substitution or swapping of high-value cuts, such as thighs for the ham industry, are presented in the *Materials and methods* section; results are discussed in the *Results and discussion* section, while *Conclusions* are addressed in the final section.

## Materials and methods

The proposed artificial intelligence (AI) method is based on a neural network that analyses the data flows from traceability data acquired along the production and delivery chains.

A supervised learning phase, based on traceability data, is carried out to provide the AI system with examples of fully compliant products. Analogously, examples of products in which one or more tracked unit variables have been perturbed along the chain (to simulate, *e.g.*, substitution, swapping, or modification) are also provided to the AI system for learning. This way, the neural network internally models discrimination rules between correct data flows and potential counterfeits without the need to be explicitly defined. This approach is particularly efficient whenever the number of ingredients or sub-products is high, and the definition of explicit, deterministic, or stochastic bounds (*i.e.*, using a strict mass balance approach) would be difficult or even impossible.

Once the neural network has been trained, it can be used to evaluate online running data flows, looking for possible data patterns which present data components that are not coherent with each other.

## Description

The case study presented in this work concerns the slaughtering and cutting phase of the pork supply chain. The proposed AI method is based on a neural network that analyses the weight data collected by the traceability system of an Italian slaughterhouse, where the meat is either sold fresh or cured.

After electrical stunning, sticking, scalding for hair and scurf removal, evisceration and splitting, the warm carcasses were weighed, classified, and then immediately delivered to the cutting line. Six meat cuts (blade, backfat, loin, ham, belly, and shoulder, as depicted in Figure 2) were selected from each carcass and weighed individually.

An identification code was assigned to both the carcass and the resulting cuts. The code was assigned to each carcass by applying a progressive number. Some cuts are subjected to controls (*e.g.*, weight and size) to verify the accordance with the strict rules of the approved PDO or IGP procedural guidelines (M.I.P.A.A.F., 2010). According to several PDO ham certification rules, if the requirements are not respected, the cut should be sold as fresh meat. Therefore, the opportunity to physically substitute one or more valuable cuts with others, which do not respect the constraints, could attract fraudsters. Thus, the link between the carcass and the cuts must not be altered to maintain the link between the carcass and the cut's quality grading data. The traceability data consist of the identification number of the carcass and its weight and the identification number of each of its cuts and weights. Since the set of cuts is only one-half of the carcass and the yield is not constant, it is impossible to apply a direct mass balance to verify the correspondence between the carcass and its cuts.

In order to implement the AI method, the supervised learning phase for the neural network training was carried out by using data acquired during the industrial slaughtering in controlled, fully compliant cutting sessions. The carcasses and main pork cuts were obtained and weighed by following the routine operating methods

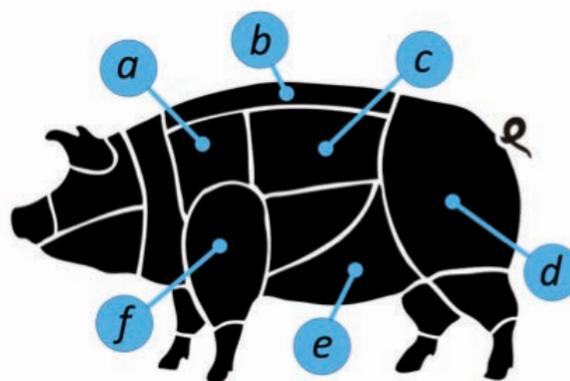


Figure 2. Pork cuts used in the neural network training phase: a- blade; b- backfat; c- loin; d- ham; e- belly; and f- shoulder.

for pork and ham processing. In order to simulate a counterfeiting event that could occur during the processing (e.g., substitution or alteration of one or more meat cuts), a further dataset of pork cut weights, obtained by perturbing the original one, was also used to train the neural network for non-compliance cases. After training, the proposed AI method can thus be implemented in advanced traceability systems to identify potential counterfeits regarding meat cuts substitution automatically.

This method may thus particularly fit in contexts of high throughput processing plants, as it would be difficult, or even impossible, to define explicit deterministic or stochastic bounds (e.g., by using a traditional mass balance approach).

### Pork cuts features

The data and the main pork cuts features were collected during a trial conducted in 2014 by the CRPA (Research Centre on Animal Production) in Reggio Emilia (Italy) and concerned a set of 134 pigs. In detail, pigs were randomly selected from a rearing farm located in northern Italy. Their breeds were 50% Goland, 36.6% Large White, and 13.4% Duroc. Using a Fat-O-Meater (FOM, an automatic device allowed by EU regulations, 2014), carcasses were classified according to their estimated lean-meat rate (kg lean meat/kg carcass) SEUROP classification (EU, 2013). The percentage of each class was 5.9% O (40-44% lean meat), 61.2% R (45-49% lean meat), 25.4% U (50-54% lean meat), and 7.5% E ( $\geq 55\%$  lean meat).

The whole group was composed of 48.5% gilt and 51.5% barrows, all 9 months old and of at least 160 kg live weight, which should be the target weight for high-quality ham production in Italy. All pigs were slaughtered at the same slaughterhouse (O.P.A.S., *Società Cooperativa Agricola, Carpi*) and processed according to industry-accepted procedures for preparing typical heavy pig carcasses. The weights of the chilled carcasses and the respective single cuts are reported in Table 1, where the means and standard deviations for the three different genetic types are also reported. A one-way analysis of variance was used to detect differences between means ( $P < 0.05$ ).

The original dataset  $A$  was classified as *compliant* as it satisfied the mandatory European requirements and the quality consortia acceptance criteria regarding the genetic type, animal weight, SEUROP carcass classification, sex, and main cuts weight. Since the objective of the neural network was to detect possible counterfeiting by product substitution, *non-compliant* data were also required to train the network properly. The *non-compliant* datasets were obtained by adopting the following data-augmentation process:

- i) dataset  $A$  was duplicated to obtain the dataset  $B$ ;
- ii) dataset  $B$  was then modified into two *non-compliant* datasets ( $C_1$  and  $C_2$ ). To do this, the data of ham weights belonging to the dataset  $B$  were randomly modified by perturbing each ham weight by  $1\sigma$  ( $\sigma$  being the standard deviation), or  $2\sigma$  thus

obtaining the  $C_1$  and  $C_2$  datasets, respectively. The  $\sigma$  value for the ham weights can be found in Table 1;

- iii) two new datasets were thus obtained as follows  $D_1 = A \cup C_1$  and  $D_2 = A \cup C_2$ ;
- iv) dataset  $D_1$  and  $D_2$  were divided in  $D'_1$  and  $D''_1$ ,  $D'_2$  and  $D''_2$ . Dataset  $D'_1$  has 85% of the data of  $D_1$ , while  $D''_1$  the remaining 15%; the same for  $D'_2$  and  $D''_2$ . Datasets  $D'_1$  and  $D'_2$  were used to train the NNs, while  $D''_1$  and  $D''_2$  were used for testing.

An additional dataset  $T$  was constructed to serve as a test set and thus simulate possible ham swapping. This could be the case when the pigs, destined to a specific supply chain (e.g., Parma ham), comply with regards to place of birth and fattening, but the hams cuts fall outside the DOP specifications. In detail, samples of datasets  $D''_1$  were swapped by substituting the ham cut weight values of the Duroc samples with samples of the Large White breed, while the Goland and Large White ham cuts were substituted with the ones of the Duroc and Goland breeds, respectively. This operation generated the  $T$  dataset that includes fraudulent swapping of ham cuts and, thus, *non-compliant* data. The two trained NNs were then used to identify possible fraud due to the substitution of original ham cuts; moreover, the method makes it possible to verify the precision of the algorithm when applied to lots of different heterogeneity levels.

### Structure of the artificial intelligence method

The algorithm used to create and train the two NNs for deep learning feature data classification was developed in the Matlab® environment (R2021a version, MathWorks®) by using a feature input layer; the two NNs were trained by using  $D'_1$  and  $D'_2$  datasets, previously described. Each dataset contains several data that consist of 8 numeric readings (weight of the whole carcass and the considered cuts) and 3 categorical labels (genetic type, sex, and FOM). To train the neural networks by using categorical features, the three categorical labels were first converted to categorical features and then to numeric ones. This conversion was automatically done using the *convertvars* and *onehotencode* functions in Matlab®.

The artificial NN structure adopted in this work for deep learning feature data classification is structured as follows:

- i) feature input layer: this layer inputs feature data into the network, applies data normalisation and sets the input size property to the specified number of features;
- ii) fully connected layer: this layer multiplies the input layer by a weight matrix and then adds a bias vector; it then returns a fully connected layer according to the specified output size property. In this application, a size of 20 was adopted. Indeed, in a fully connected layer, all neurons connect to all the neurons in the previous layer;
- iii) batch normalisation layer: this layer normalises a mini-batch of data across all observations for each channel independently.

**Table 1. Mean weights and standard deviation of the chilled carcasses and the respective single cuts (ham, shoulder, blade, loin, belly, and backfat).**

Breed	Cold carcass (kg)	Blade (kg)	Shoulder (kg)	Ham (kg)	Loin (kg)	Belly (kg)	Backfat (kg)
Duroc	130.1±9.5 <sup>a</sup>	4.7±0.4 <sup>a</sup>	9.6±0.8 <sup>a</sup>	17.2±1.3 <sup>a</sup>	12.6±1.1 <sup>a</sup>	9.0±1.1 <sup>a</sup>	4.0±0.9 <sup>a</sup>
Goland	135.1±9.4 <sup>a</sup>	5.1±0.5 <sup>a</sup>	10.1±0.8 <sup>a</sup>	17.7±1.4 <sup>a</sup>	12.9±1.1 <sup>a</sup>	9.5±1.5 <sup>a</sup>	3.9±1.0 <sup>a</sup>
Large White	133.0±11.9 <sup>a</sup>	4.8±0.5 <sup>a</sup>	9.9±0.9 <sup>a</sup>	18.0±1.8 <sup>a</sup>	12.9±1.2 <sup>a</sup>	9.1±1.5 <sup>a</sup>	3.8±1.0 <sup>a</sup>
Total	133.7±10.4	4.9±0.5	10.0±0.9	17.8±1.5	12.9±1.1	9.3±1.4	3.9±0.9

Means with the same superscript letters are not significantly different.

- After normalisation, to avoid those inputs with zero mean affecting the normalised data activation, the layer scales and shifts the input using two parameters updated during network training;
- iv) rectified linear unit (ReLU) layer: a ReLU layer is a nonlinear activation function that performs a threshold operation on each input element, where any value less than zero is set to zero. The ReLU layer does not change the size of its input;
  - v) fully connected layer: a second fully connected layer, with its output size corresponding to the number of classes, is used for data classification;
  - vi) softmax layer: this layer applies a softmax activation function to the input; such function is also known as the normalised exponential and can be considered the multi-class generalisation of the logistics sigmoid function;
  - vii) classification output layer: this layer usually follows the softmax layer and classifies the values obtained by the softmax function by assigning each input to one of the classes using the cross-entropy function.

The *trainNetwork* Matlab® function was used to train the network employing the architecture defined by the 7 layers described, the training and validation data, and the training options (e.g., tolerance). Datasets  $D'_1$  and  $D'_2$  were partitioned into training and validation subsets by using an array of random indices; in detail, 70% of the data of  $D'_1$  and  $D'_2$  were used for training and 30% for validation to avoid overfitting. During training, a standard gradient descent algorithm based on an adaptive moment estimation (Adam optimiser) was used. The NNs were validated at regular intervals during the training phase. The validation data were not used to update the network weights nor for testing. Datasets  $D''_1$  and  $D''_2$  were used for testing the trained NNs. The  $T$  dataset was used only for testing the performance of the NNs, i.e., to test the ability of the neural networks to detect carcasses and meat cuts in which some substitutions have been carried out.

The Matlab® algorithm was run on a MacBook Pro 13" which has a 2.3 GHz i7 quad-core processor and 32 GB 3733 MHz of RAM.

## Results and discussion

The neural networks were trained using Matlab®, and the convergence of each neural network training and validation phase was obtained by running the algorithm for thirty epochs according to

**Table 2. Confusion matrix of the neural network trained by using the  $D'_1$  dataset (85% of  $D_1$ ). The number of samples is lower than the  $D_1$  cardinality as the only subset  $D''_1$  (15% of  $D_1$ ) was used for testing.**

		Predicted class	
		Compliant	Non-compliant
True class	Compliant	17 (42.1%) - TP	3 (7.9%) - FN
	Non-compliant	2 (5.3%) - FP	18 (44.7%) - TN
TPR =		$\frac{TP}{TP+FN} = 85.0\%$	
Precision =		$\frac{TP}{TP+FP} = 89.5\%$	
FPR =		$\frac{FP}{FP+TN} = 10.0\%$	

TP, true positive; FN, false negative; FP, false positive; TN, true negative; TPR, true positive rate; FPR, false positive rate.

the adopted tolerance. The average elapsed time to train and validate the neural networks was 9 seconds; this result mainly depends on the size of the dataset in terms of the number of samples (rows), the number of features (columns), and data quality.

The obtained accuracy of the trained neural networks for the two datasets  $D'_1$  and  $D'_2$  was 86.8% and 100%, respectively. A confusion matrix based on  $D''_1$  and  $D''_2$  datasets was calculated for each neural network, and the results were reported in Tables 2 and 3 in terms of true positive (TP), false positive (FP), false negative (FN), and true negative (TN). Tables 2 and 3 also report: i) the true positive rate (TPR), which is the measure of how many true positives get predicted out of all the positives in the dataset; ii) the precision, which is a measure of the correctness of a positive prediction; and iii) the false positive rate (FPR), which is the measure of how many results get predicted as positive out of all the negative cases. The TPR and precision values obtained for  $D''_2$  are higher than those obtained for  $D''_1$ ; as might be expected, the FPR values decrease with increasing TPR values. Samples of the  $T$  dataset were also used for testing the neural network trained with the  $D'_1$  dataset; in detail, about 86% of the altered samples of dataset  $T$  were correctly identified.

These results show the ability of the AI method to detect pattern anomalies at different levels of the input perturbation. Furthermore, since the neural networks were trained by using different kinds of information (numeric features such as cold carcass weight, single cuts weight, as well as categorical labels such as sex, breed, etc.), this method is more robust than standard mass balance methods since it can simultaneously analyse numeric readings, statistics, and categorical labels. This can be assumed to be representative of the skill of the AI system in detecting weight variations in single pig cuts, which could be related to food fraud.

The findings prove that artificial intelligence methods can be used to implement new features in traditional food traceability as they can monitor and check the reliability and the truthfulness of the data flow in a concise time. This characteristic is essential as the promptness of supervision systems in fraud identification is essential to apply measures for the repression of illicit actions.

In particular, the AI method adopted in this work can be applied to strengthen traceability in critical traceability points. To apply the method, the neural network must be properly trained on data obtained by considering pooled pig samples, choosing among those types usually processed in the plant(s). Furthermore, this methodology can be further improved in terms of robustness to

**Table 3. Confusion matrix of the neural network trained by using the  $D'_2$  dataset (85% of  $D_2$ ). The number of samples is lower than the  $D_2$  cardinality as the only subset  $D''_2$  (15% of  $D_2$ ) was used for testing.**

		Predicted class	
		Compliant	Non-compliant
True class	Compliant	20 (50.0%) - TP	0 (0%) - FN
	Non-compliant	0 (0%) - FP	20 (50.0%) - TN
TPR =		$\frac{TP}{TP+FN} = 100\%$	
Precision =		$\frac{TP}{TP+FP} = 100\%$	
FPR =		$\frac{FP}{FP+TN} = 0\%$	

TP, true positive; FN, false negative; FP, false positive; TN, true negative; TPR, true positive rate; FPR, false positive rate.

possible fraud attempts, adopting blockchain technologies to guarantee the inviolability of food traceability information in terms of data and transactions.

## Conclusions

This paper presents an innovative artificial intelligence method for online traceability data analysis aimed at identifying possible counterfeiting events. This method could be exploited in systems that generate a warning whenever a possible inconsistency in the traceability data flow is detected. Since this method is based on direct learning from the data, it does not require any explicit pre-definition of rules, thresholds or statistical discriminators. The reported example shows how this method can easily be implemented, is computationally efficient, and is able to detect the swapping of single items (meat cuts). The opportunity to consider numerical and categorical variables can significantly enhance the number of applications of neural networks in food traceability. This approach, which considers mass balance data congruity as a tool to detect possible fraud associated with product substitution or modification, is not based on data integrity protection from cyber intrusions, which can be obtained using other techniques (e.g., blockchain). Nevertheless, this method can be recommended when enforcing traceability systems at the points of the production/delivery chain where lots are split or joined or, more generally, wherever the granularity is changed. Artificial intelligence methods embedded in traceability systems could also be favourably adopted to identify the occurrence of other types of fraud by selecting and monitoring the appropriate production process variables.

## References

- Al-Hashedi K.G., Magalingam P. 2021. Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Comput. Sci. Rev.* 40:100402.
- Barge P., Biglia A., Comba L., Gay P., Ricauda Aimonino D., Tortia C. 2017. Temperature and position effect on readability of passive UHF RFID labels for beverage packaging. *Chem. Engine. Trans.* 58:169-74.
- Barge P., Biglia A., Comba L., Gay P., Ricauda Aimonino D., Tortia C. 2019. The influence of food composition and tag orientation on UHF Rf IDentification. *J. Food Eng.* 246:242-52.
- Barge P., Biglia A., Comba L., Ricauda Aimonino D., Tortia C., Gay P. 2020. Radio Frequency IDentification for meat supply chain digitalisation. *Sensors* 20:4957.
- Böhme K., Calo-Mata P., Barros-Velázquez J., Ortea I. 2019. Review of recent DNA-based methods for main food-authenticity topics. *J. Agric. Food Chem.* 67:3854-64.
- Comba L., Belforte G., Gay P. 2011. Modeling techniques for the control of thermal exchanges in mixed continuous-discontinuous flow food plants. *J. Food Eng.* 106:177-87.
- Comba L., Belforte G., Dabbene F., Gay P. 2013. Methods for traceability in food production processes involving bulk products. *Biosyst. Eng.* 116:51-63.
- Dabbene F., Gay P., Tortia C. 2014. Traceability issues in food supply chain management: a review. *Biosyst. Eng.* 120:65-80.
- Dey S., Saha S., Singh A.K., McDonald-Maier K. 2021. FoodSQRBlock: Digitizing food production and the supply chain with blockchain and QR code in the cloud. *Sustainability* 13:3486.
- Dupuy C., Botta-Genoulaz V., Guinet A. 2005. Batch dispersion model to optimize traceability in food industry. *J. Food Eng.* 70:333-9.
- El Sheikha A.F. 2021. 1 - Food authentication: Introduction, techniques, and prospects. *Food Authent. Trace.* 1-34.
- Esteki M., Cardador M.J., Jurado-Campos N., Martín-Gómez A., Arce L., Simal-Gandara J. 2021. Chapter 8 - Innovations in analytical methods for food authenticity. *Innov. Food Analysis* 181-248.
- European Commission. 1996. Regulation (EC) No 1107/96 of 12 June 1996 on the registration of geographical indications and designations of origin under the procedure laid down in Article 17 of Council Regulation (EEC) No 2081/92. *Official Journal L* 148, 21/06/1996 P. 0001-0010.
- European Commission. 2009. Regulation 1239/2009 of 15 December 2009 entering a name in the register of protected designations of origin and protected geographical indications Crudo di Cuneo (PDO). *GU L* 332, 17.12.2009, 50-51.
- European Commission. 2014. Commission implementation decision of 24 January 2014 authorising methods for grading pig carcasses in Italy, *Official Journal of the European Union* 28/1/2014, 57, 35-40.
- European Union, December 2013. Regulation (EU) No 1308/2013 of the European Parliament and of the Council establishing a common organisation of the markets in agricultural products and repealing Council Regulations (EEC) No 922/72, (EEC) No 234/79, (EC) No 1037/2001 and (EC) No 1234/2007.
- Galimberti A., De Mattia F., Losa A., Bruni I., Federici S., Casiraghi M., et al. 2013. DNA barcoding as a new tool for food traceability. *Food Res. Int.* 50:55-63.
- Hellberg R., Everstine K., Sklare S. 2021. A Global threat with public health and economic consequences. *Food Fraud.* doi:10.1016/B978-0-12-817242-1.00021-X.
- Jahanbakhshi A., Abbaspour-Gilandeh Y., Heidarbeigi K., Momeny M. 2021. A novel method based on machine vision system and deep learning to detect fraud in turmeric powder. *Comput. Biol. Med.* 136:104728.
- Karlsen K.M., Donnelly K.A.M., Olsen P. 2011. Granularity and its importance for traceability in a farmed salmon supply chain. *J. Food Eng.* 102:1-8.
- Kemény Zs., Ilie-Zudor E. 2016. 4 - Alphanumerical and optical coding systems for food traceability. *Adv. Food Trace. Technol.* 49-65.
- Kendall H., Clark B., Rhymer C., Kuznesof S., Hajslova J., Tomaniova M. 2019. A systematic review of consumer perceptions of food fraud and authenticity: A European perspective. *Trends Food Sci. Tech.* 94:79-90.
- MathWorks® 2021. Available from: <https://it.mathworks.com/products/matlab.html>
- M.I.P.A.A.F. - Ministero delle Politiche Agricole e Alimentari della Repubblica Italiana 2010. Provvedimento 1 marzo 2010 Modifica del disciplinare di produzione della denominazione «Prosciutto di Parma» registrata in qualità di denominazione di origine protetta in forza al regolamento (CE) 1107 della Commissione del 12 giugno 1996. (10A02925). *Gazzetta Ufficiale Serie Generale*, 66 del 20-03-2010, 1-80.
- Mol A.P.J., Oosterveer P. 2015. Certification of markets, markets of certificates: Tracing sustainability in global agro-food value chains. *Sustainability* 7:12258-78.
- Niknejad N., Ismail W., Bahari M., Hendradi R., Salleh A.Z. 2021. Mapping the research trends on blockchain technology in food and agriculture industry: A bibliometric analysis. *Environ. Technol. Innov.* 21:101272.

- Robson K., Dean M., Haughey S., Elliot C. 2021. A comprehensive review of food fraud terminologies and food fraud mitigation guides. *Food Control* 120:107516.
- Soon J.M., Manning L. 2019. Developing anti-counterfeiting measures: The role of smart packaging. *Food Res. Int.* 123:135-43.
- Spink J.W., Ortega D.L., Chen C., Wu F. 2017. Food fraud prevention shifts the food risk focus to vulnerability. *Trends Food Sci. Tech.* 62:215-20.
- Spink J.W. 2019. The current state of food fraud prevention: overview and requirements to address 'How to Start?' and 'How Much is Enough?'. *Curr. Opin. Food Sci.* 27:130-8.
- Spink J.W., Chen W., Zhang G., Speier-Pero C. 2019. Introducing the Food Fraud Prevention Cycle (FFPC): A dynamic information management and strategic roadmap. *Food Control* 105:233-41.
- Vo S.A., Scanlan J., Turner P. 2020. An application of convolutional neural network to lobster grading in the southern rock lobster supply chain. *Food Control* 113:107184.
- Wadood S.A., Boli G., Xiaowen Z., Hussain I., Yimin W. 2020. Recent development in the application of analytical techniques for the traceability and authenticity of food of plant origin. *Microchem. J.* 152:104295.
- Wang J., Yue H., Zhou Z. 2017. An improved traceability system for food quality assurance and evaluation based on fuzzy classification and neural network. *Food Control* 79:363-70.
- Wilkinson S., Archibald A.L., Haley C.S., Megens H.J., Crooijmans R.P.M.A., Groenen M.A.M., 2012. Development of a genetic tool for product regulation in the diverse British pig breed market. *BMC Genomics* 13:580.
- Xu J., Guo S, Xie D., Yan Y. 2020. Blockchain: A new safeguard for agri-foods. *Artif. Intell. Agric.* 4:153-61.
- Young M., Himmelreich J., Honcharov D., Soundarajan S. 2021. The right tool for the job? Assessing the use of artificial intelligence for identifying administrative errors. *DG.O 2021: The 22nd Annu. Int. Conf. Digital Govern. Res.* 15-26.

Non-commercial use only