International Journal of Food Science +Technology

Institute of Food Science +Technology ifst

Original article

Prediction of fish quality level with machine learning

Emre Yavuzer[1]* **& Memduh Köse**[2]

1 Department of Food Engineering, Faculty of Engineering and Architecture, Kırşehir Ahi Evran University, 40100, Kırşehir, Turkey
2 Department of Electrical Electronics Engineering, Faculty of Engineering and Architecture, Kırşehir Ahi Evran University, 40100, Kırşehir, Turkey

(Received 9 April 2022; Accepted in revised form 19 May 2022)

Summary In this study, sea bream, sea bass, anchovy and trout were captured and recorded using a digital camera during refrigerated storage for 7 days. In addition, their total viable counts (TVC) were determined on a daily basis. Based on the TVC, each fish was classified as 'fresh' when it was <5 log cfu per g, and as 'not fresh' when it was >7 log cfu per g. They were uploaded on a web-based machine learning software called Teachable Machine (TM), which was trained about the pupils and heads of the fish. In addition, images of each species from different angles were uploaded to the software in order to ensure the recognition of fish species by TM. The data of the study indicated that the TM was able to distinguish fish species with high accuracy rates and achieved over 86% success in estimating the freshness of the fish species tested.

Keywords Food identification, fresh fish, machine learning, quality changes, teachable machine.

Introduction

Machine learning refers to the artificial intelligence consisting of various algorithms, which can process the information uploaded to the system, turn them into useful information and use them for making estimations. Machine learning has initiated important changes in many industries in recent years and has enabled the elimination of important problems thanks to its high performance. The major advantage of machine learning is its ability to process large amounts of data rapidly with a time-saving checkpoint (Xu *et al.*, 2021). In recent years, learning machines have acquired the capacity to process the images provided to them, particularly enabling the effective use of facial recognition modules in many locations such as hospitals, banks, schools, security cameras or workplaces (Visweswaraiah *et al.*, 2017; Sharma *et al.*, 2021). Especially after the emergence of the COVID-19 pandemic, different machine learning systems have been developed, where contactless tracking techniques are utilised (Sharma *et al.*, 2021).

Machine learning can be used for determining food quality as an essential method that can be applied without contacting and destroying food. In relevant previous studies (Igathinathane *et al.*, 2009; Moallem *et al.*, 2017), significant results were obtained, especially on foods such as apples, which can be examined for food quality using machines. In addition, machine learning has been used in the measurement of parameters such as the amount of toxin in foods (Wu *et al.*, 2018), cracked egg shells (Deng *et al.*, 2010), degree of milling in rice (Zareiforoush *et al.*, 2015) and colour of food (Minz *et al.*, 2020).

With regard to preservation of fish as a functional food, it is desirable that fatty acids and proteins can be delivered to the consumer without being destroyed. Fish is rich in polyunsaturated fatty acids (Özoğul *et al.*, 2013; Yavuzer, 2018; Roy *et al.*, 2019), which makes it a perishable food that spoil quickly. While the susceptibility of aquatic foods to oxidative degradation is one of the most important problems in terms of product storage in processing technology (Yavuzer, 2020), undesired changes in the taste, odour and colour of the product due to oxidative deterioration lead to a decrease in their shelf life (Singh *et al.*, 2005). Therefore, lipid peroxidation and microbial contamination, which are considered as the main causes of deterioration (Cao *et al.*, 2008), cause significant changes in the colour, shape and appearance of the fish. Chemical, sensory and microbiological quality analyses are usually performed for monitoring these changes.

The first version of the Teachable Machine (TM) was launched by Google in 2017. It is a fast, easy and accessible web-based tool, which creates machine learning models. TM uses the TensorFlow.js library, which is used for machine learning in Javascript, to

*Correspondent: E-mail: emre.yavuzer@ahievran.edu.tr

doi:10.1111/ijfs.15853

train and run models created in the web environment and enables to develop individual deep learning models without any coding. In TM, the image files uploaded from a camera and the images available on the computer are quickly added into the database. In this study, the main reasons for using TM in determining the parameters of fish quality include the usability of the software by anyone without knowledge of coding, its ability to save the project and run on other computers, and easy selection of the learning criteria such as epoch, batch size or learning rate. In the study, TM was trained on the head and eye structures of four different fish species, which varied during the storage period. The fact that TM was able to interpret the data obtained enabled the determination of the parameters regarding the freshness of the fish quickly by the machine.

Materials and methods

Machine learning program

The 'Teachable Learning' platform, which does not require special expertise or technical knowledge, was used in the study. Epoch: 50, Batch size: 16 and learning rate 0.001 settings were used in the training of the program. Fifty different fish appearances were used for fresh and spoiled fish entering the system.

Deep learning

Deep learning networks can be employed for classification purposes directly using the data itself without feature extraction phase. The methods based on deep learning networks may be performed on a wide range of data types including one or two-dimensional signals. In general, large data sets are required in these methods in order to achieve high performance. Convolutional neural networks (CNN) are the most well-known deep learning networks. In CNN, input data are convolved with the learned features and 2D convolutional layers are employed (Mathworks, 2022a). This feature of CNN makes it suitable for image processing. In image classification applications, training of CNN is carried out by using a set of images rather than using the manually extracted features.

In this study, a support vector machine (SVM) algorithm is used to classify the images in which CNN features obtained from images were employed. In this approach, features were extracted by using a pretrained CNN which reduces effort in feature extraction stage (Mathworks, 2022b). As a CNN, ResNet-50 model defined in MATLAB Deep Learning Toolbox was used. The first layer of CNN specifies the input dimensions of the images. When the images have different dimensions, resizing functions can be employed to satisfy the size requirements. The intermediate layers following the input layer are the series of rectified linear units, maximum pooling layers, interspersed convolution layers (Krizhevsky *et al.,* 2012). These layers are connected to the classification layer where features are classified.

Fish material and imaging conditions

Rainbow trout (*Oncorhynchus mykiss*), sea bream (*Sparus aurata*), sea bass (*Dicentrarchus labrax*) and anchovy (*Engraulis encrasicolus*) were obtained from a local fish market located in Kırşehir/Turkey. All groups of fish were stored in a refrigerator (3 ± 1 °C) in polystyrene boxes for 7 days. Fish were selected randomly (just by reaching into the box without looking) and all images were taken in daylight conditions with a 64-megapixel phone camera. The category in which the image will be used (fresh or spoiled) was determined based on the total viable count (TVC). In the study, it was determined that loading TM head together with eye data increased the accuracy estimates in all groups. Figure S1 illustrates the eye and head data of fresh seabass.

Microbiological analysis

Total viable counts were made using the petri dish smear method (ICMSF, 1986). Triplicate samples were taken for TVC. Ninety ml of sterile Ringer solution (1/4 strength) was mixed with 10 g of fish muscle and then stomached (Masticator Nr S18/420, IUL Instruments, Barcelona, Spain) for 3 min. More decimal dilutions were made, and then 0.1 mL of each dilution was pipetted onto the surface of plate count agar (Fluka 70 152, Steinheim, Switzerland) plates in triplicate. After that, plates were incubated for 2 days at 30 °C.

Results and discussion

Figure S2 displays the TVC occurring in the fish meat of trout, sea bream, sea bass and anchovy during 7 days of storage. Images, which were taken on the storage days displaying values smaller than 5 log cfu per g, were used to train TM about fresh fish. The storage days, in which the fish were determined as 'fresh', were until 3 days for trout, sea bream and sea bass; and 2 days for anchovy. Groups above the level of 7 log cfu per g were used in the process of creating a database for the 'spoiled fish'. While trout reached this level on the 5th day of storage, sea bream and sea bass reached on the 4th day, and anchovy reached on the 3rd day. Bensid *et al.* (2014) found the initial total mesophilic bacteria count of anchovy was 4.33 log cfu per g and reached 5.82 log cfu per g on the 3rd day of storage. In the same study, TBA level, which had been 3.08 mg MA per kg at the beginning of storage, was found as 8.54 mg MA per kg on the third day of storage and as 14.38 mg MA per kg on the sixth day of storage. Similar to present study, these data indicate that the anchovy is considered as fresh on the 2nd day of storage. Parlapani *et al.* (2015) found the initial TVC value of sea bass as 3 log cfu per g, which increased during storage. In another study conducted on sea bream (Parlapani *et al.,* 2014), the initial TVC value was found as 3.9 log cfu per g, and the bacterial load at the time the fish lost its freshness was reported as 7.5–8.5 log cfu per g. In addition, it was determined that trout started to spoil when the bacterial load exceeded the limit of 7 log cfu per g (Yavuzer, 2020; Yavuzer *et al.,* 2020; Yavuzer, 2021).

Figure S3 displays the estimation for species identification performed by TM which was trained for sea bream, sea bass, trout and anchovy on the images of the same species randomly selected from the Internet. The accuracy rate of the estimations of sea bream among the randomly selected images were recorded between 95 and 100%. This value ranged between 89 and 100% in the sea bass group and 75 and 100% in anchovy group. Trout was the group that was the most confused with sea bass by TM, and the accuracy rate of the estimations on trout varied between 55 and 100%. While the images of trout and sea bass on the internet were generally found similar by TM, when the real fish were presented to TM instead of the images, the accuracy rate of the software was 80% and above in all test groups.

Table 1 presents the changes in eye structures of four fish groups during the 7 of day storage period. The eye structures of the fish groups analysed in the study were bright at the beginning of storage, and they became cloudy towards the end of the storage, as expected. During storage, while the halos around the pupils of the sea bream gradually turned white; the eye halos of the sea bass, which used to be black when fresh, turned white. The eyes of anchovy, which was the group with the fastest-changing eye structure, started to become cloudy even on the 2nd day of storage when it was <3 log cfu per g, compared with the first day of storage.

In the current study, the freshness level of each group was determined separately on the basis of species. For this purpose, TM was trained about fish heads and pupils captures on storage days in order to ensure the quality estimation of the whole fish. During the quality estimation, the eyes of the fish tested were clear reflected in a way that the TM could learn, which increased the success of estimation. Figure S4 displays the estimation performed by TM regarding the quality parameters of sea bass.

In order for TM to distinguish between fresh and spoiled fish properly, the images were taken in a standard light and background. Accordingly, as the quality of the images increased, the accurate estimation rate of TM also increased. The training of TM and the increase in the rate of accurate estimation were directly proportional to the number of data loaded. For example, while the number of pupils uploaded for fresh and spoiled fish was 10 in the sea bream group (Figure S5), the accurate estimation rate of TM was found to be below 40%. When the number of pupils uploaded was increased to 50, the accuracy rate of TM reached over 80%.

Despite the small number of data uploaded, the group with the highest estimation rate was anchovy among all tested groups. Figure S6 displays one of the estimations performed by TM for the quality of anchovy. In the study, anchovies, which were not really in good shape in terms of appearance and microbiologically, which started to smell and had blood in the eyes, were estimated accurately at a quite high rate (>90).

In the present study, TM was very successful in estimating fish species and quality. Nowadays, the ability of many mobile phones to take high-quality images makes it easier to create the database required for image processing. It can be concluded that these software can be used at least in the preliminary inspection process of food quality control, since software such as TM can process the image without the need for coding skills and have a high rate of accurate estimation. In addition, the TM may also be recommended to be used in the determination of spinal curvatures such as scoliosis, eye defects and fungal infections on the skin, which are diagnosed defectively during the processing of fish, as well as the determination of fish diseases under aquaculture conditions.

The data set contains images captured from four different fish species within seven days. The images captured on the first two days were labelled as ‘fresh’ and the images captured on the following five days were labelled as ‘not fresh’. In Table 2, the number of images in each category is presented.

In this study, two classification tasks were performed. In the first classification task, images from each species were classified as ‘fresh’ or ‘not fresh’. In the second classification task, images were classified with respect to fish species. For both the classification tasks, data set was divided into training and test data sets such that each of these sets contain 50% of the total images included in the data set. In order to perform Monte Carlo cross-validation, 100 trials were performed, and average classification accuracy was calculated. Randomly selected training and test sets were used in each trial. The training and test sets were processed in the CNN to obtain features, then

Table 1 Changes in eye structures of fish groups during 7 days of storage

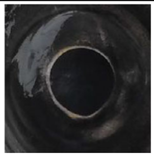
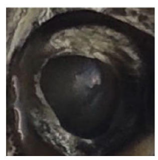
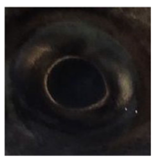
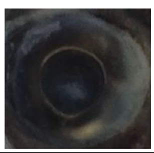
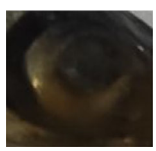
Storage days	Sea bream	Sea bass	Anchovy	Trout
1				
2				
3				
4				
5				
6				
7				

Table 2 Number of fish images in each category in the data set

	Fresh	Not fresh	Total
Trout	46	136	182
Sea bream	86	100	186
Anchovy	79	91	170
Sea bass	82	87	169

SVM algorithm was employed to classify the features as explained in deep learning subsection. The confusion matrix containing correct classification and misclassification rates was calculated for each trial. Then, the mean confusion matrix over 100 trials was calculated.

The mean confusion matrix for fresh/not fresh classification is given in Table S1. The average classification accuracy was calculated by averaging the diagonal elements of the confusion matrix in Table S1. The average correct classification rates for trout, sea bream, anchovy and sea bass were obtained as 86%, 94%, 98% and 97%, respectively and given at the top of each table. In the fresh/not fresh classification, the best classification performance was obtained for anchovy with a rate of 98%, and the lowest classification performance result was obtained for trout with a rate of 86%.

The confusion matrix of fish species classification is given in Table S2. Correct classification rates vary between 90.6% and 99.8%. The highest classification accuracy was obtained for anchovy with a rate of 99.8%. The lowest classification performance was obtained for sea bream, in which case sea bream was misclassified as sea bass and trout with ratios of 6.25% and 2.05%, respectively.

Conclusion

In this study, a web-based image processing software was used to quickly determine the freshness parameters of fish. Besides, classification was performed by using a CNN-based algorithm. The obtained results show that from these two approaches perform similarly. The results of the study indicated that the determination of the fish species, and distinction between fresh and spoiled fish can be performed with high accuracy using image processing. Uploading the learning codes on TM with a single click and ensuring its availability on operating systems such as Android will enable the determination of fish quality with mobile devices such as mobile phones. The effective performance of TM- and CNN-based algorithm in determining fish quality will inspire future studies to be conducted on fruit and vegetable quality or diseases.

Ethical approval

Ethics approval was not required for this research.

Author contributions

Emre Yavuzer: Conceptualization (lead); data curation (equal); formal analysis (equal); funding acquisition (lead); investigation (lead); methodology (equal); project administration (lead); resources (lead); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Memduh Köse:** Data curation (equal); formal analysis (equal); methodology (equal); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal).

Peer review

The peer review history for this article is available at https://publons.com/publon/10.1111/ijfs.15853.

Data availability statement

The data that supports the finding of this study are available in the supplementary material of this article.

References

Bensid, A., Uçar, Y., Bendeddouche, B. & Özoğul, F. (2014). Effect of the icing with thyme, oregano and clove extracts on quality parameters of gutted and beheaded anchovy (*Engraulis encrasicholus*) during chilled storage. *Food Chemistry*, **145**, 681–686.

Cao, L., Si, J.Y., Liu, Y. *et al.* (2008). Essential oil composition, antimicrobial and antioxidant properties of Mosla chinensis Maxim. *Food Chemistry*, **115**, 801–805.

Deng, X., Wang, Q., Chen, H. & Xie, H. (2010). Eggshell crack detection using a wavelet-based support vector machine. *Computers and electronics in agriculture*, **70**, 135–143.

ICMSF (1986). Microorganisms in Foods. In: *The International Commission on Microbiological Specifications for Foods of the International Union of Biological Societies* Pp. 181–196. Oxford: Blackwell Scientific Publications

Igathinathane, C., Pordesimo, L.O. & Batchelor, W.D. (2009). Major orthogonal dimensions measurement of food grains by machine vision using ImageJ. *Food Research International*, **42**, 76–84.

Krizhevsky, A., Sutskever, I. & Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12)* Pp. 1097–1105. Red Hook, NY, USA: Curran Associates Inc

Mathworks,(2022a). *Deep Learning*. Retrieved from https://www.mathworks.com/discovery/deep-learning.html Accessed March 30, 2022.

Mathworks, (2022b). *Image Category Classification Using Deep Learning*. Retrieved from https://www.mathworks.com/help/vision/ug/image-category-classification-using-deep-learning.html Accessed March 28, 2022.

Minz, P.S., Sawhney, I.K. & Saini, C.S. (2020). Algorithm for processing high definition images for food colourimetry. *Measurement*, **158**, 107670

Moallem, P., Serajoddin, A. & Pourghassem, H. (2017). Computer vision-based apple grading for golden delicious apples based on surface features. *Information Processing in Agriculture*, **4**, 33–40.

Özoğul, F., Yavuzer, E., Özoğul, Y. & Kuley, E. (2013). Comparative quality loss in wild and cultured rainbow trout (*Oncorhynchus mykiss*) during chilling storage. *Food Science and Technology Research*, **19**, 445–454.

Parlapani, F.F., Haroutounian, S.A., Nychas, G.J.E. & Boziaris, I.S. (2015). Microbiological spoilage and volatiles production of gutted European sea bass stored under air and commercial modified atmosphere package at 2 °C. *Food Microbiology*, **50**, 44–53.

Parlapani, F.F., Mallouchos, A., Haroutounian, S.A. & Boziaris, I.S. (2014). Microbiological spoilage and investigation of volatile profile during storage of sea bream fillets under various conditions. *International Journal of Food Microbiology*, **189**, 153–163.

Roy, J., Mercier, Y., Tonnet, L. *et al.* (2019). Rainbow trout prefer diets rich in omega-3 long chain polyunsaturated fatty acids DHA and EPA. *Physiology & Behavior*, **213**, 112692

Sharma, N., Sharma, R. & Jindal, N. (2021). Machine learning and deep learning applications-A vision. *Global Transitions Proceedings*, **2**, 24–28.

Singh, G., Marimuthu, P., Murali, H.S. & Bawa, A.S. (2005). Antoxidative and antibacterial potentials of essential oils and extracts isolated from various spice materials. *Journal of Food Safety*, **25**, 130–145.

Visweswaraiah, M., Somashekar, K. & Babu, N.V. (2017). Test mode power computation and IR drop analysis of application specific integrated circuits implementing face detection algorithms. In: *4th International Conference on Advanced Computing and Communication Systems (ICACCS)*. Institute of Electrical and Electronic Engineering, pp. 1–4.

Wu, Q., Xie, L. & Xu, H. (2018). Determination of toxigenic fungi and aflatoxins in nuts and dried fruits using imaging and spectroscopic techniques. *Food Chemistry*, **252**, 228–242.

Xu, Y., Zhou, Y., Sekula, P. & Ding, L. (2021). Machine learning in construction: From shallow to deep learning. *Developments in the Built Environment*, **6**, 100045

Yavuzer, E. (2018). Development of defective fish egg sorting machine with colour sensor for trout facilities. *Aquaculture Research*, **49**, 3634–3637.

Yavuzer, E. (2020). Determination of rainbow trout quality parameters with Arduino microcontroller. *Journal of Food Safety*, **40**, 1–7.

Yavuzer, E. (2021). Determination of fish quality parameters with low cost electronic nose. *Food Bioscience*, **41**, 100948

Yavuzer, E., Özogul, F. & Özogul, Y. (2020). Impact of icing with potato, sweet potato, sugar beet, and red beet peel extract on the sensory, chemical, and microbiological changes of rainbow trout (*Oncorhynchus mykiss*) fillets stored at (3 ± 1 °C). *Aquaculture International*, **28**, 187–197.

Zareiforoush, H., Minaei, S., Alizadeh, M.R. & Banakar, A. (2015). A hybrid intelligent approach based on computer vision and fuzzy logic for quality measurement of milled rice. *Measurement*, **66**, 26–34.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Data set consisting of fresh sea bass images.

Figure S2. Changes in the total viable count of fish species during storage.

Figure S3. Species prediction by TM of sea bream (a), sea bass (b), anchovy (c) and trout (d).

Figure S4. Estimation of freshness level of sea bass by TM.

Figure S5. Estimation of freshness level of sea bream by TM.

Figure S6. Estimation of freshness level of anchovy by TM.

Table S1. Fresh/not fresh classification confusion matrices in each category.

Table S2. Confusion matrix for fish species classification.