Fish Freshness Estimation though analysis of Multispectral Images with Convolutional Neural Networks

G. Tsagkatakis[*], S. Nikolidakis[†], E. Petra[‡], A. Kapantagakis[†], K. Grigorakis[†], G. Katselis[⋆], N. Vlahos[⋆], P. Tsakalides[*]
[*] ***Institute of Computer Science, Foundataion for Reseach and Technology - Hellas (FORTH), 70013, Heraklion, Crete, Greece***
[†] ***Hellenic Centre for Marine Research, 19013, Anavyssos, Attiki, Greece***
[‡] ***Information Management Systems Institute, Athena Research Center, 15125, Marousi, Greece***
[⋆] ***Department of Fisheries and Aquiculture Technology, University of Patras, 30200, Messolonghi, Greece***

Abstract

Quantification of food quality is a critical process for ensuring public health. Fish correspond to a particularly challenging case due to its high perishable nature as food. Existing approaches require laboratory testing, a laborious and time-consuming process. In this paper, we propose a novel approach for evaluating fish freshness by exploiting the information encoded in the spectral profile acquired by a snapshot spectral camera. To extract the relevant information, we employ state-of-the-art Convolutional Neural Networks and treat the problem as an instance of multi-class classification, where each class corresponds to a two-day period since harvesting. Experimental evaluation on individuals from the Sparidae (Boops sp.) family demonstrates that the proposed approach constitutes a valid methodology, offering both accuracy as well as effortless application.

Introduction

Fish are the most vulnerable foods since changes in fish fleshness are very rapid due to bacterial, enzymatic and oxidative causes. Fish spoilage is influenced by several factors such as the type of fish, the way of extinction and the method of preservation. Freshness, in addition to the quality value it offers to fish as a food, has a significant impact on its commercial value as a marketable commodity. European legislation categorizes fish into four categories based on their quality value as food, namely Excellent, Category A, Category B and Inadmissible. In the first category, the catches are classified immediately after their catch and in the latter category unsuitable for human consumption [1].

The traditional approach for estimating the freshness of a dietary product is using human senses like appearance, odor, flavor and texture [2]. Human senses however exhibit a very high degree of subjectivity and therefore are questioned in cases of a dispute. To address this issue, numerous laboratory evaluation processes have been developed. In this work, we describe an approach for estimating fish freshness by analyzing the spectral profile obtained from multispectral imagery using state-of-the-art machine learning methods and particularly, Convolutional Neural Networks (CNNs). The proposed approach offers a number of significant benefits compared to existing approaches including the extremely fast evaluation, the ability to perform in-situ estimation, and the non-discrictive nature of the method. The novelties of this work include:

- The use of a state-of-the-art CNN for the automated extraction of appropriate spatio-spectral features;
- Use of multispectral information for a snapshot spectral camera;
- The creation of a new dataset which is made publicly available;
- The demonstration of fish freshness estimation system using optical information.

The objective of this paper is the demonstration of the capabilities of state-of-the-art deep learning architectures and more specifically Convolutional Neural Networks (CNN) in analyzing multispectral images for estimating the freshness of fish. By fish freshness, we mean the estimation of the apparent days since harvest, which may not coincide with the actual number of days in cases of problematic storage conditions. The targeted scenario involves a compact and portable system which will consist of a snapshot spectral imaging system and a mini-pc system like the NVIDIA Jetson, while the interfacing with the user will be achieved though transmission of the processed images to his/her smartphone, as shown in Figure 1.

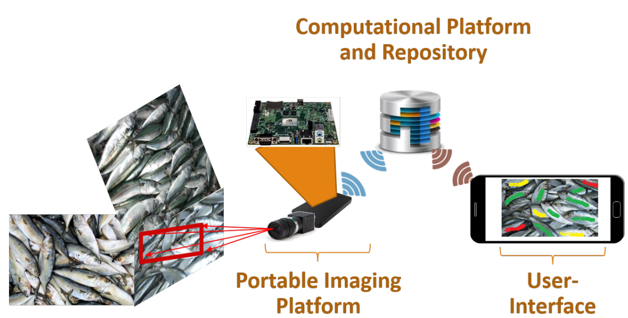

Figure 1. *Overview of the system.*

State-of-the-art

Degradation in fish quality is due to the posthumous enzyme activity which leads to the breakdown of cell membranes which in turn leads to the exponential increase in the bacterial load, causing the sever degradation of the skin. Currently, different types of instruments are available for the quantification of fish freshness which exploit the alteration of different physiological parameters. Different approaches that have been presented in the liturate include methods based on enzyme biosensors, electrochemical biosensors, colorimetric sensor, electronic tongue, and different types of spectroscopy[3]. Multispectral imaging, a particular case of spectroscopy is among the most prominent solutions for this problem due to the numerous benefits such method offer including the in-situ and real-time estimation. In this work, we employ Snapshot Spectral Imaging technologies which in addition to the general benefits of spectroscopy, do not require scaning the item over a conveyor belt. This techology has been recently employed for classifying generic objects [6], as well for food quality monitoring including detecting plant desease [7] and red-meat classification [8].

Method

The proposed system currently consists of a snapshot spectral imaging camera and a data processing pipeline based on machine learning for estimating the elapsed time between harvesting and imaging. Specifically, observations are acquired over the visible-near IR range (400-1000nm) using a snapshot spectral camera from Photon Focus (MV1-D2048x1088-HS02-96-G2), equipped with an IMEC sensors acquiring snapshot spectral imagery of 2048×1088 pixels at 42fps. Unlike traditional linescan approaches, is able to acquire the entire spectral profile of a scene from a single image (exposure). This capability makes snapshot spectral cameras ideal for the deployment in real-time environments, removing the need for specialized platforms like conveyor platforms. The quantum efficiency of each band is shown in Figure 2.

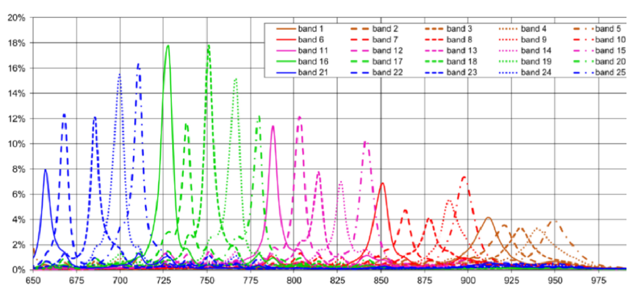

Figure 2. *Quantum efficiency associated with each spectral band.*

An example of the appearance of fish individuals at two different wavelength on day 1 and day 4 is shown in Figure 3. In this example, one can observe that regions around the tail and around the eye appear to change appearance with the passing of time, an observation which is consistent with the known sensory tests.

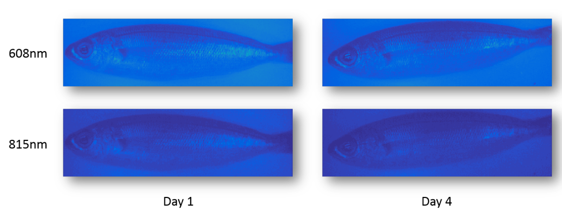

Figure 3. *Examples of fish imaged at* 608 *and* 815*nm on the first and forth day.*

The acquired image is subsequently introducing into a Convolutional Neural Network (CNN) which consist of four convolutional layers and one fully connected layer. The specifics of the considered architecture is shown in Figure 5. A significant difference between the proposed architecture and typical architectures employed in image classification is that the spatial size of the image is rectangular and that the multispectral images consist of 25 spectral bands, much higher compared to 3 in color imagery.

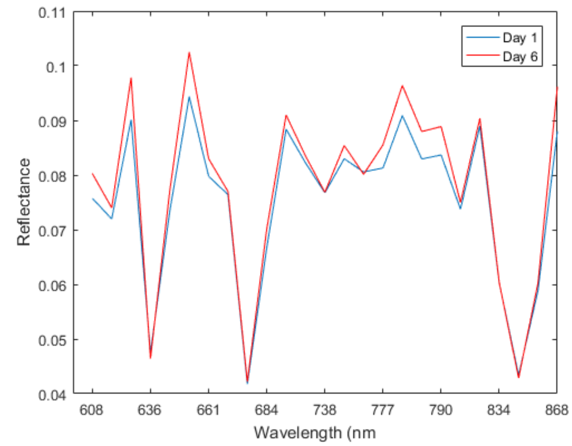

Figure 4. *Reflectance at different wavelengths for an individual fish at days 1 and 6 (selected region around the eye).*

Layer (type)	Output Shape	#Parameters
Input Layer	(2, 120, 409, 25, 1)	0
Conv3D	(2, 120, 409, 25, 8)	152
Batch Normalization	(2, 120, 409, 25, 8)	32
Max Pooling 3D	(2, 60, 205, 13, 8)	0
Conv3D	(2, 60, 205, 13, 8)	1160
Batch Normalization	(2, 60, 205, 13, 8)	32
Max Pooling 3D	(2, 30, 103, 7, 8)	0
Conv3D	(2, 30, 103, 7, 8)	1160
Batch Normalization	(2, 30, 103, 7, 8)	32
Max Pooling 3D	(2, 15, 52, 4, 8)	0
Conv3D	(2, 15, 52, 4, 8)	1160
Batch Normalization	(2, 15, 52, 4, 8)	32
Max Pooling 3D	(2, 8, 26, 2, 8)	0
Flatten	(2, 3328)	0
Dropout	(2, 3328)	0
Dense	(2, 6)	19974
Total params: 23,734		

Figure 5. *Network architecture including dimensions for each layer and associated number of parameters.*

A significant novelty of the proposed framework is that freshness is quantified in terms of days since harvest and thus corresponds to a continuous variable. In order to fully exploit the capabilities of the CNN and provide high accuracy prediction, despite the limited number of training examples, the output is discretized, reformulated the problem from a regression to a multi-class classification framework.

Typically, in multi-class classification problems, the loss function which is networks is trained to minimize is the categorical cross entropy. This function is optimal when each outputs class is independent from each other. In our case however, this is not the case, since the error between the true class and the predicted class corresponds to temporal distance between the true and the predicted harvested days. In order to encode this information in our loss function, we employ the 1D Wasserstein distance [] where the output of the network, as well as the ground truth, are treated as a probability density functions, assigning a probability to each class. Formally, given distributions u and v and the associated CDF's U and V, the 1D Wasserstein distance between u, v is given by

$$D(u,v) = \int |U-V|^p \approx \sum_i |U-V|^p \tag{1}$$

In this work, we consider the ℓ_1 metric by setting $p = 1$.

Experimental results

We evaluate the method on fish from the Sparidae family (Boops boops) family and perform daily acquisition of multispectral images from 5 individuals for a period of 6 days, leading to a total of 30 examples. In order to train and validate the performance of the system, 24 examples were utilized for training and 6 for testing. We consider two cases, prediction at a single and a two two days resolution, leading to either a 3 or a 6 class classification problems from a time frame of 6 days between harvesting and measurement.

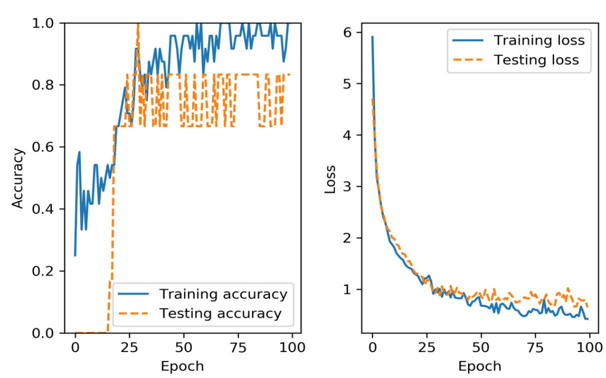

Figure 6. Accuracy and loss function value for training and testing sets for the **3** *class problem.*

Figure 6 showcases the accuracy and loss associated with training and validation examples for a network trained over 100 epochs for the case of two day resolution, i.e. 3 classes. The results indicate that the network is able to achieve perfect (100%) accuracy on the training data, while for the evaluation data, the accuracy is at 83%. Examination of the progression of the value of the loss function, indicates that the difference in accuracy between training and validation can most likely be attributed to a case of

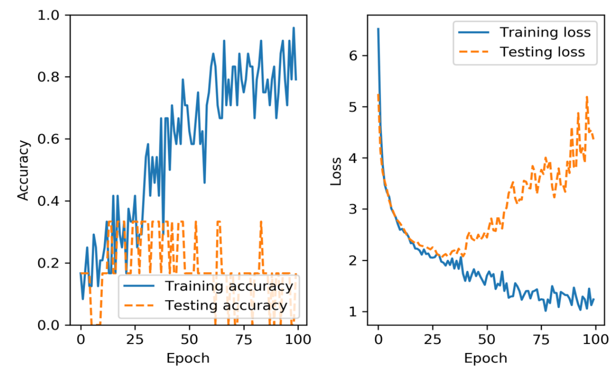

Figure 7. Accuracy and loss function value for training and testing sets for the **6** *class problem.*

overfitting, so which the most effect solution is the use of a larger number of training examples.

The impact of overfitting is even more pronounced in the case of single day resolution, i.e., 6 class classification, as demonstrated in terms of accuracy and loss value in Figure 7. In this case, although the accuracy for the training set reaches 79%, for the case of the validation examples, the best reported accuracy is 16%, while the value of the loss function between training and validation sets begins to rapidly diverse from as early at 25 epochs.

An example of features extracted from different convolution layers along the network hierarchy is shown in Figure 8. In this figure, one can observe that features extracted from deeper layers tend to focus on regions around the eye and the tail. This observation is in-line with physics based model and sensory method which indicate that indeed these regions are the most important for judging the freshness of fish.

Figures 6 and 7 present the confusion matrices for the testing data for the cases of 3 and 6 classes. An observation which can be made from these confusion matrices is that even for the cases where the network is confused, the predicted differs from the truth by a single day.

Conclusions

In this paper we present our initial results on the utilization of multispectral imagery for the estimation of fish freshness using a deep learning algorithm. The results indicate that the system can achieve very high prediction accuracy on coarse time scales while for finer time scale, it suffers from overfitting. Nevertheless, even for misclassification cases, the prediction differs from the ground truth by a single day, an error which for this application can be considered acceptable. A major limitation of this work is the limited amount of examples available for both training and testing the learning system. Future work will focus on collecting a larger set of examples, which will hopefully also address the overfitting issues.

Acknowledgments

This work was funded by the CALCHAS project contract no.842560 within the H2020 Framework Program of the European Commission.

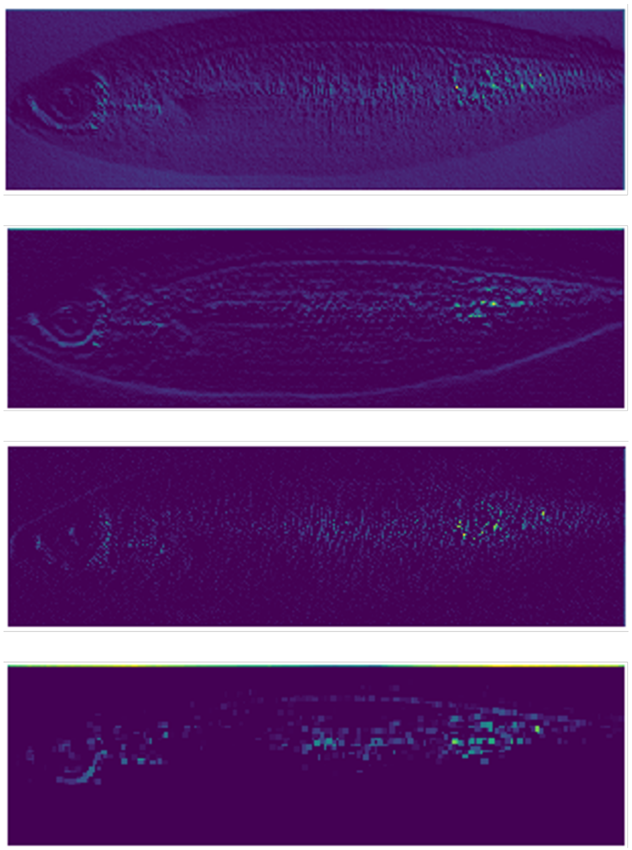

Figure 8. Feature maps extracted from 1st, 2nd, 3rd and 4th convolutional layer for a given test image.

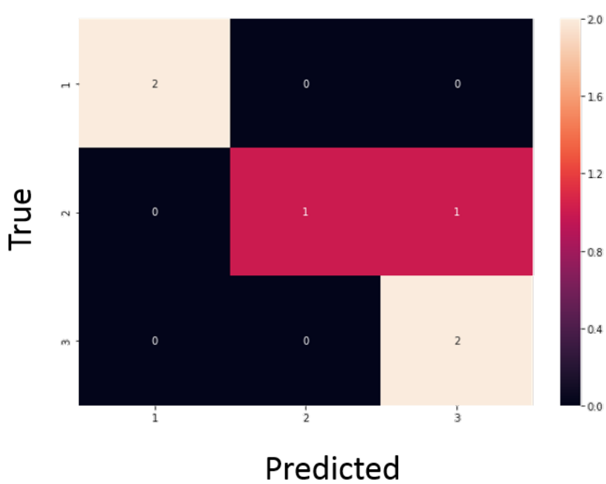

Figure 9. Confusion matrix for the case of **3** classes.

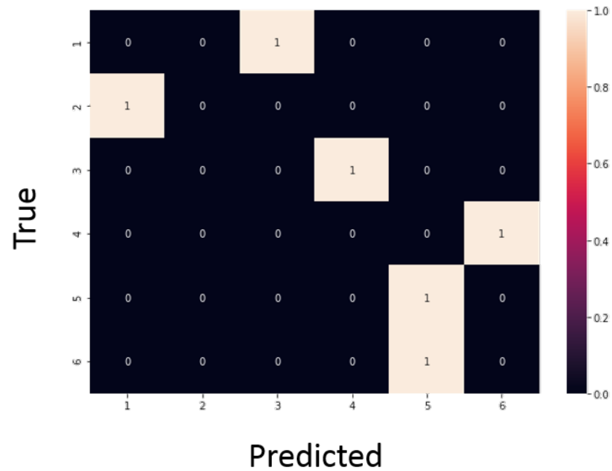

Figure 10. Confusion matrix for the case of **6** classes.

References

[1] Cheng, J.H., Sun, D.W., Han, Z. and Zeng, X.A., 2014. Texture and structure measurements and analyses for evaluation of fish and fillet freshness quality: a review. Comprehensive Reviews in Food Science and Food Safety, 13(1), pp.52-61.

[2] Martinsdóttir, E., Sveinsdottir, K., Luten, J., Schelvis-Smit, R. and Hyldig, G., 2001. Sensory evaluation of fish freshness. IJmuiden, The Netherlands: QIM eurofish.

[3] Wu, L., Pu, H. and Sun, D.W., 2019. Novel techniques for evaluating freshness quality attributes of fish: A review of recent developments. Trends in food science & technology, 83, pp.259-273.

[4] LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. nature, 521(7553), pp.436-444.

[5] Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[6] Fotiadou, K., Tsagkatakis, G. and Tsakalides, P., 2017. Deep convolutional neural networks for the classification of snapshot mosaic hyperspectral imagery. Electronic Imaging, 2017(17), pp.185-190.

[7] Nagasubramanian, K., Jones, S., Singh, A.K., Singh, A., Ganapathysubramanian, B. and Sarkar, S., 2018. Explaining hyperspectral imaging based plant disease identification: 3D CNN and saliency maps. arXiv preprint arXiv:1804.08831.

[8] Al-Sarayreh, M., Reis, M.M., Yan, W.Q. and Klette, R., 2018, November. Deep spectral-spatial features of snapshot hyperspectral images for red-meat classification. In 2018 International Conference on Image and Vision Computing New Zealand (IVCNZ) (pp. 1-6). IEEE.