Prediction fat percentage and visceral weight from whole fish images with a multi-input neural network

Y. Xue[1*], J.W.M. Bastiaansen[1] and H. Komen[1]

[1] Wageningen University & Research, P.O. Box 338, 6700 AH Wageningen, the Netherlands; [*]yuanxu.xue@wur.nl

Abstract
In aquaculture, high accuracy in trait measurements benefits the genetic progress from a breeding program. Breeding traits like fat percentage and visceral weight are related to feed/cost efficiency of growth and product quality, and important metabolism and health indicators. Problems concentrate on finding the proper methods to accurately measure or predict these traits, as most current approaches are invasive, labour-intensive or may disturb or damage the fish. Interior trait prediction from image analysis would allow a real-time, large-scale and non-invasive alternative for such traits. This study investigates using whole-fish images in combination with exterior traits to improve the prediction of fillet fat percentage and visceral weight. The result of including images as extra input shows improvement on the accuracy of fat percentage prediction. The neural network extracted contour-based features and brings into view several biological indicators that appear to be informative for prediction.

Introduction
In aquaculture, accurate measurements of breeding traits on individuals are essential for the assessment of breeding values and the progress of the breeding program. Modern breeding programs emphasize not only on production traits, but also traits that are important indicators of metabolism and health of the fish, such as fat percentage and visceral weight. However, these two traits are costly to measure for two reasons: the required measurement involves stress challenges and may require sacrificing the fish, risking the loss of a potential breeding candidate; these measurements are invasive and damaging, raising societal and ethical concerns (Seibel, Weirup & Schulz, 2020).

To answer the need for accurate assessment of these traits, researchers and breeders start to include indirect measurements. Popular approaches include fitting prediction models using easy-to-measure exterior traits to approximate the fat percentage and visceral weights (García-Celdrán *et al*., 2015, Vandeputte *et al*., 2017). Another approach, image analysis, has received increasing attention in recent years (Prchal *et al*., 2020). However, image analysis often heavily relies on morphological annotation. This transfer of image data to interpretable numeric value inevitably discards other image features that are potentially informative to prediction. Color pattern, for instance, is difficult to be objectively quantified as a morphological reading. Yet the information is easy to obtain non-invasively and can be valuable to assess the health status of fish (Parolini *et al*., 2018). Combined with the proper statistical methods, images might offer more than annotated predictors for invasive traits.

This study estimates the accuracy of predicting fat percentage and visceral weight by different models. It aims to elucidate if adding whole-fish image as extra inputs is beneficial to predict these two traits, and if so, how images contribute to the prediction.

Materials & Methods
Data was collected on a total of 4766 gilthead seabream (*Sparus aurata*) and kept separated in two groups. Images for these seabreams were taken with a Logitech HD Pro Webcam C920.

Group A. Records of fat percentage and body weights on 3069 individuals with their corresponding images. Data was modelled for the predictability of body fat percentage based on body weight with and without images.

Group B. Records of visceral weights, fat percentage and body weights on 1697 individuals with their corresponding images. Data was modelled for the predictability of visceral weight based on body weight, body fat percentage, and with and without images.

For each group, data was divided into training, validation and test sets with a size ratio of 3:1:1. Training and validation sets were used for exploratory data analysis, cross validation and hyperparameter tuning. To avoid information leakage, test set was only used once for the final evaluation of the prediction. Three methods were included for comparison on their prediction performance.

Linear regression (LR). Although the assumed relationship between the variables and the response is simple, linear models are highly robustness with low computing cost. Given its simplicity and robustness, LR is chosen as the benchmark to use body weight for prediction of fat percentage and to use body weight and fat percentage for prediction of visceral weight. LR was combined with 5-fold cross validation to avoid overfitting.

Multilayer perceptron (MLP). Trait prediction sometimes requires more complex model as the relationship between the predictors and response are not always linear. Therefore, this study uses MLP to represent a simplified collection of all complex prediction model like polynomial regression, sinusoidal regression, etc. It is intended to capture the real relationship by going through all possible combinations of mathematical connections between body weight, fat percentage, and visceral weight to use body weight for prediction of fat percentage and to use body weight and fat percentage for prediction of visceral weight. Images are not included as predictors.
MLP is in essence a collection of neurons. Each neuron consists of a value and relates to either an input or an output neuron, or both. Neurons are grouped per layer. The connections between layers form a chain of mathematical functions that continuously manipulate the original input towards a final output. These functions can be as simple as a subtraction or addition or as sophisticated as a derivative. At the end of every round of calculation, these functions will self-update to minimize the difference between the calculated and observed output until the difference reaches the minimum. This is the learning process through which a MLP finally realizes a prediction.

Multi-input neural network (NN). To incorporate images to the prediction, a multi-input NN is applied. It is meant to include all pixel-level information from a whole-fish image, regardless of the biological interpretability of the information, and combine it with other measurement like body weights. A multi-input NN consists of two parts: a convolutional neural network (CNN) and an MLP. CNN scans the images repeatedly using different filters. These filters form different layers like neurons in MLP and translate the value of raw image pixels into imaginal features and concepts. In parallel, the MLP establishes a mathematical relationship between the numeric input from other traits like body weight. The outputs of these two compartments are combined and matched with the response variables. Multi-input NN uses the same 'learning

process' as MLP to update all connected layers, until the difference between the observed output and the prediction stabilizes and reaches its minimum.
Feature visualization helps to visualize the learning process by displaying the filtered features from each convolutional layer. Mathematically, these features maximize the information that contributes to the observed output. According to the human visual comprehension, these features can vary from edges and textures to patterns and parts of the objects.

The predictions were evaluated using two metrics: Mean squared error (MSE) and Pearson correlation coefficient (*r*). Lower MSE and higher *r* values imply that a prediction model has better prediction performance.

Results
Table 1 shows the results of prediction fitting three models on both test set. Accuracy of predicting visceral weight is high and similar for all methods (0.87 – 0.90). Prediction of fat percentage benefits from the image data which increases accuracy from 0.48 to 0.70.

Table 1. Evaluation of prediction methods on test sets.

	Group A - Fat percentage		Group B - Visceral weight	
Metrics	MSE[1]	r	MSE[2]	r
LR	4.86	0.48	63.26	0.90
MLP	5.26	0.48	75.41	0.89
Multi-Input NN (with images)	2.98	0.70	87.14	0.87

[1] Test set has a mean of 11.52 and a standard deviation of 2.5
[2] Test set has a mean of 40 and a standard deviation of 18

To understand how images facilitate the prediction, the extracted features by each convolutional layer are displayed in figure 1. Each row contains features obtained from the same layer. Moving to the next row, the layer goes deeper, which means that the features extracted will become more abstract but also more informative to the final prediction. The first row in figure 1 illustrates that in the beginning layer, multi-input NN extracts different features, like whole-fish contours and contrasts between regions, as informative inputs for fat percentage prediction. Colour might also be considered as an informative feature given the visual inconsistency in the background. Visually dark regions on the fish correspond to locations like patch, eye and fin on the colour image. However, the multi-input NN quickly abandons this information and fixates on outlines and contours, starting from the second row of figure 1. The final convolutional layer in the last row comprises features mostly highlighted in visually dark colour: the curvature of the dorsal side, together with the upper edge of pectoral fins.

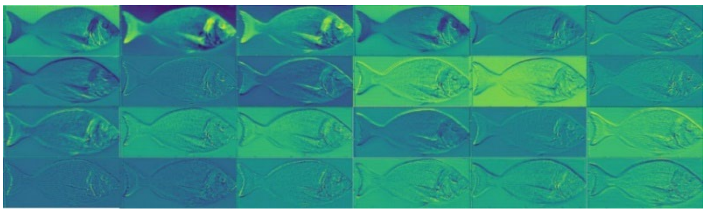

Figure 1. Visualized extracted features by each convolutional layer of the multi-input neural network from Group A, fat percentage.

Figure 2 compares the final convolutional features in both groups. Like its performance in group A on fat percentage, multi-input NN applied to visceral weights in group B has a strong preference for the contour. The right image in figure 2 shows an overlapping highlighted pattern around the edges, indicating that the contour has been read repetitively for visceral weight prediction. Compared to the result in fat percentage, however, there are additional patterns revealed on the body of the fish, especially around the operculum, dorsal edge and anal fin.

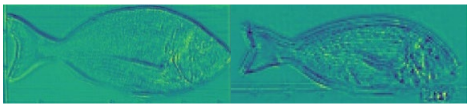

Figure 2. Visualized extracted features from the last convolutional layer of the multi-input neural networks. Left: Group A, fat percentage; Right: Group B, visceral weight.

Discussion

For fat percentage prediction, accuracy of LR and MLP are similar. This similarity on accuracy indicates that the non-linear, complex MLP failed to capture more information compared to a linear model for predicting fat percentage. In contrast, multi-input NN uses the images and outperformed both LR and MLP. Features extracted suggest two important factors in predicting fat percentage: contour and pectoral fin. Contour can be interpreted as the size of body area, which is strongly correlated with body weight (Fernandes *et al*., 2020). However, given the poor performance of LR with only body weight as predictor, the multi-input NN probably uses a different interpretation of the contour. Curvature is hypothesized to be the informative factor in fat percentage prediction. The result of MLP implies that prediction accuracy does not benefit from mathematically complex indicators. Including the images as inputs identifies curvature which is also a rather explicit morphological description. Further work could separate the contour and investigate a proper quantitative description of the curvature and of the location of pectoral fin and test the predictive value of these features. However, without a quantitative description of the contour, images can still improve the fat percentage prediction by the feature extraction of multi-input NN.

For visceral weight prediction, three models yielded similar results, suggesting that neither the complexity of the MLP method nor the information-concentrated image are good additions for visceral weight prediction. Future work may look into replacing body weight by body features detected from images as a fast, non-invasive alternative to assess the visceral weight.

Acknowledgement

This project receives financial support from the Topsector Agri & Food.

References

García-Celdrán, M., Ramis, G., Manchado, M., Estévez, A., Afonso, J. M. *et al*. (2015). Aquaculture 446:175–180. https://doi.org/10.1016/j.aquaculture.2015.04.028

Fernandes, A. F., Turra, E. M., de Alvarenga, É. R., Passafaro, T. L., Lopes, F. B. *et al*. (2020). Computers and electronics in agriculture, 170, 105274. https://doi.org/10.1016/j.compag.2020.105274

Parolini, M., Iacobuzio, R., Bassano, B., Pennati, R., and Saino, N. (2018). Physiological and Biochemical Zoology, 91(5): 1026–1035. https://doi.org/10.1086/699522

Prchal, M., Kocour, M., Vandeputte, M., Kause, A., Vergnet, A. et al. (2020) Aquaculture, 520. https://doi.org/10.1016/j.aquaculture.2020.734993

Seibel, H., Weirup, L., and Schulz, C. (2020). Food Ethics, (1–2). https://doi.org/10.1007/s41055-019-00063-3.

Vandeputte, M., Puledda, A., Tyran, A. S., Bestin, A., Coulombet, C. *et al*. (2017) Aquaculture 470:40–49. https://doi.org/10.1016/j.aquaculture.2016.12.014