Deep learning based fish length estimation. An application in the Mediterranean aquaculture

Dimitris Voskakis[1], Alexandros Makris[2] and Nikos Papandroulakis[1]

1. *Hellenic Center for Marine Research, I.M.B.B.C., Crete, Greece*
2. *Foundation of Research and Technology Hellas, I.C.S., Crete Greece*

e-mail: d.voskakis@hcmr.gr

Aquaculture farms have become recently a pole of interest, with innovative technologies gradually substituting labor intensive procedures. An effective farm management requires accurate biomass estimation and estimates of the individual growth performance during the grow out (Beddow et al., 1996). For individual growth estimation, current practice involves on site sampling which is both labor intensive and stressful for the fish. In this work, we present a non-invasive system of individual fish size estimation that relies on stereoscopy and machine vision/artificial intelligence algorithms.

There exist optical methods tackling the problem of biomass estimation applying stereoscopy (e.g (Harvey et al., 2003) for individual measurements while there are some exceptions that fish body length is estimated by one camera (Monkman et al., 2020). The proposed method uses convolutional neural networks (CNN) which have proven successful for the task (Cao et al., 2019).

The system consists of two IP cameras located in a submersible housing. A mini computer handles the frame acquisition and synchronization, and uploads the data to an off-site server via ethernet. Camera calibration is carried out using a chessboard pattern to correct lens and water distortions. The system acquires the synchronized and calibrated images from both cameras (Image acquisition) and using a trained CNN detects key points of individual fish (snout, eye, pelvic fin and fork tail). As the 2D skeleton (pose) of the same fish in both images are detected, the two dimensional skeleton coordinates are subsequently translated into the three dimensional space coordinates by correlating the corresponding points of the same scene in both cameras based on stereoscopy (Redon, n.d.). Given the 3D coordinates the standard length of each fish is estimated. To train the keypoint detection network we created a dataset of manually annotated fish images. The process is presented in Figure 1.

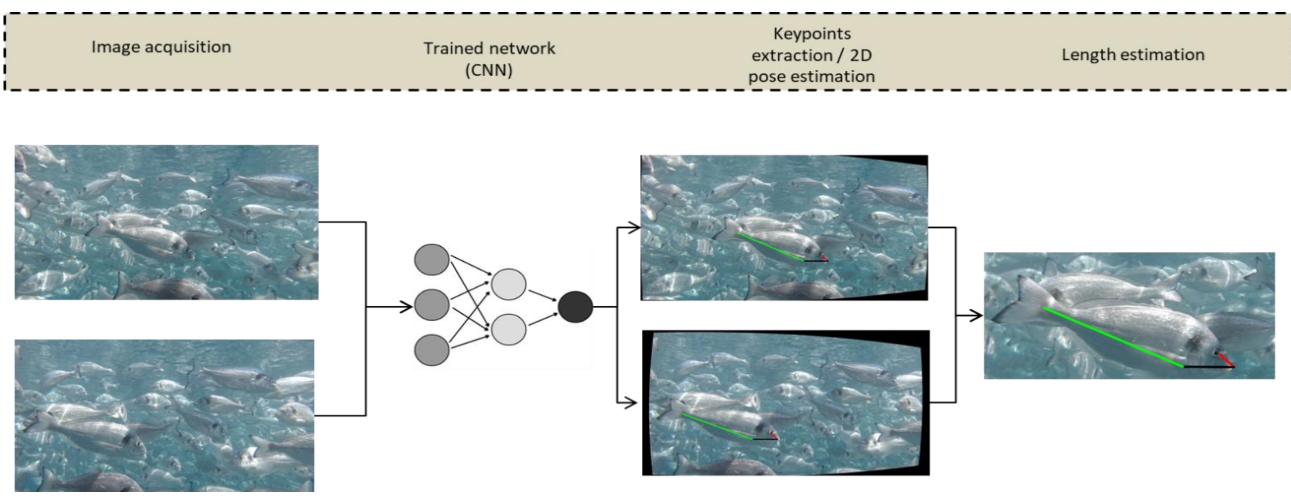

Figure 1. A set of synchronized images is given as an input in a trained network and the 2D skeleton (pose) of the same fish in both frames is detected. Afterwards, the two-dimensional to three-dimensional points translation is performed using stereoscopy allowing us to compute the snout and fork length. Subsequently, the standard length is calculated.

The system was tested at the pilot netpen aquaculture farm of HCMR in Crete, Greece. Initially, the calibration process was performed to estimate the intrinsic and extrinsic camera parameters. Using the calibrated cameras we collected image data of the two main Mediterranean aquaculture species, namely the gilthead sea bream (*Sparus aurata*) and European sea bass (*Dicentrarchus labrax*). A total of apx 2,000 images per species were used. The system was tested on calm and windy days, where the light intensity was variable. The results were compared with values (mean length and variability) taken applying standard sampling methods mean length.

The system resulted in high accuracy estimations of the standard length for both species , with mean relative error of 3.15% and 7.4%, for the gilthead sea bream and the European sea bass respectively. Main source of error was the environmental conditions during the sampling that, due to waves and the increase of suspended matter in the water column, added an “optical noise” to the collected images. Fish movements were another major factor causing errors due to self-overlapping. In fact, the introduction of the system itself in the cage during the sampling resulted in a behavioral reaction of the fish that in some cases inhibited the process until were habituated with its presence. The system could be further refined, making it more robust to difficult weather conditions and trained to detect more species of interest for the aquaculture industry.

Acknowledgements

This study has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 727610 (PerformFISH). This output reflects only the authors’ views and the European Union cannot be held responsible for any use that may be made of the information contained therein.

References

Beddow, T.A., Ross, L.G., Marchant, J.A., 1996. Predicting salmon biomass remotely using a digital stereo-imaging technique. Aquaculture 146, 189–203. https://doi.org/10.1016/S0044-8486(96)01384-1

Harvey, E., Cappo, M., Shortis, M., Robson, S., Buchanan, J., Speare, P., 2003. The accuracy and precision of underwater measurements of length and maximum body depth of southern bluefin tuna (Thunnus maccoyii) with a stereo–video camera system. Fisheries Research 63, 315–326. https://doi.org/10.1016/S0165-7836(03)00080-8

Redon, O., n.d. Camera calibration and epipolar geometry 57.

Monkman, G.G., Hyder, K., Kaiser, M.J., Vidal, F.P., 2020. Accurate estimation of fish length in single camera photogrammetry with a fiducial marker. ICES Journal of Marine Science 77, 2245–2254. https://doi.org/10.1093/icesjms/fsz030