A Large-Scale Fish Dataset Statistics, Analysis and Prediction

By:

Harsh Raj
Manipal University Jaipur
Dehmi Kalan , Off Jaipur-Ajmer
Expressway, Jaipur, Rajasthan 303007,India,
harsh.219302137@muj.manipal.edu

***Abstract**: It is critical to assess the freshness of fish in retail settings and throughout the manufacturing process to avoid spoilage, which can lead to serious health problems and financial losses. A key component of this strategy is evidence of species recognition, as many fish species exhibit distinct symptoms of illness and deterioration. Traditional approaches to species classification can be tedious and sleepy in some situations, so more persuasive methods are required.*

Image processing and machine learning frameworks appear to be dependable solutions that deliver precise results quickly. In any case, developing these systems necessitates the collection of relevant and interesting datasets. Regrettably, a large number of freely available datasets—many of which include fish types that are not commonly consumed or include submerged images—are contaminated.

Inappropriate for our purposes. This study fills that void by developing a valuable and comprehensive dataset that includes nine common fish species that are commonly consumed in Turkey's Aegean region.

The investigation also conducts broad tests utilizing distinct categorization procedures to evaluate the dataset's utility. The tests' highly promising results demonstrate the dataset's sufficiency. As a result, this data will be made available to the public to support further research in the area of fish quality assessment.

Keywords—Fish dataset, feature extraction, segmentation, classification, food quality assessment

I. Introduction

Seafood plays a crucial role in various dishes, particularly in coastal countries, due to its essence and nutritional value. The number of fish consumed annually worldwide has surpassed 20 kilograms, which has made it more difficult for the food business to satisfy customer demand for premium items on short notice. To satisfy customers and prevent financial losses, seafood must be kept fresh. Labor-intensive and time-consuming are the traditional ways of identifying spoilage, such as laboratory analysis. Furthermore, subjective decisions and the unintentional sale of damaged seafood might result from depending too much on subjective assessments, which are frequently based on visual examination. The creation of an automated spoilage detection system that is dependable, rapid, and objective is essential to overcoming these obstacles. Before separating fresh fish from rotten ones and recognizing illnesses in different species, automated classification is required since different fish breeds display different signs as they acquire diseases and spoil.

Fish texture, size, and form are frequently used as features in feature-based fish categorization systems. For example, research utilizing support vector machines (SVMs)-based classifiers and combining ear bone morphological data with sex, length, and weight variables obtained up to 75% classification accuracy. An important challenge in fish classification studies is the lack of publicly available datasets containing commonly consumed fish image samples. Already present datasets contain fish images taken underwater and consist of fish species that are not usually consumed [10]–[15]. These datasets mostly provide data for marine biologists, scientists, and researchers; hence they are not practically suitable for the food quality assessment problem. To the best of our knowledge, no publicly available dataset contains seafood and fish sold in retail. This study fulfils the need for such a dataset containing image samples of eight different fish species and shrimp which are collected from the fish counter of a supermarket. Also, comprehensive classification tasks are performed to analyse the usability of this dataset. In detail, the paper provides analysis results of the dataset through semantic segmentation and feature-based classification relying on grey-level cooccurrence matrices (GLCM), moments, bag-of-features (BoF), and CNNs features (CNN) via SVMs. The detailed analysis

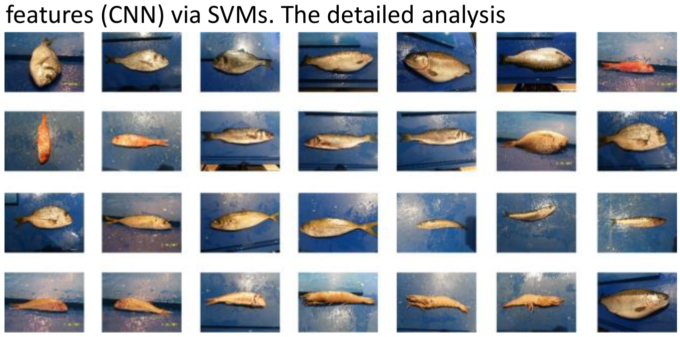


Figure 1: Example images from the collected dataset.

of this dataset will provide a guide light for future research, as it demonstrates success rates of the employed methods in seafood segmentation and classification tasks. Furthermore, the collected dataset and its ground-truth (manually extracted) segmentation masks will be publicly available for research purposes, which is one of the main contributions of this study. The rest of the paper is organized as follows. Section II presents the details of the collected dataset, and then demonstrates the experimental setup and results. Section III concludes this study with a summary.

II. EXPERIMENTAL SETUP AND RESULTS

A. The Dataset

The dataset used in this study contains images from nine different angle types taken with two cameras (, Kodak Easy Share Z650, and Samsung ST60). The spatial size of the image is 2832 x 2128 pixels or 1024 x 768 pixels. The seven classes include his 50 images of fish species such as red mullet, red mullet, horse mackerel, perch, sea bream, black porgy, and striped mullet. Additionally, he added 30 images of trout and shrimp. Despite the different introductions and angles, the illumination conditions remain consistently similar across the dataset Notably, the dataset deviates from the traditional pure white background and opts for a blue, vibrant undercoat to mimic real-world environments. A test image from the captured dataset is shown in Figure 1.To standardize the dataset, all images of the nine classes are resized to 590 x 445 pixels while maintaining their original aspect ratio. An extended procedure involving non-iterative random rotations and reflections is used. This process creates a final batch of 1,000 images for each angle type, further increasing the diversity and richness of the dataset. Before delving into SVM-based classification, the analysis includes regression and Z-test techniques in a test setup.

Regression analysis examines relationships between variables and allows you to explore possible correlations within a data set. This step provides insight into how specific features are related and influence the overall classification performance.

After regression analysis, use the Z-test to assess the significance of the observed differences.

The Z-test helps determine whether observed variation within a dataset is statistically significant or simply due to chance.

This evaluation of statistical significance is important for determining the reliability of the dataset and subsequent classification results.

Figure 2 shows the comprehensive test setup including regression, Z-test analysis, and SVM-based classification.

These preliminary steps serve as a basis for evaluating the robustness of the dataset and the effectiveness of the classification strategy under different feature types.

In the following sections, we present the results of these analyses and highlight the effectiveness of our dataset and the performance of our classification strategies, with a particular focus on SVM and different feature types.

This comprehensive approach provides valuable insight into the strength and relevance of datasets in real-world scenarios.

B. Fish Segmentation

In the first set of experiments, the ground-truth segmentation masks of all seafood images are extracted manually by

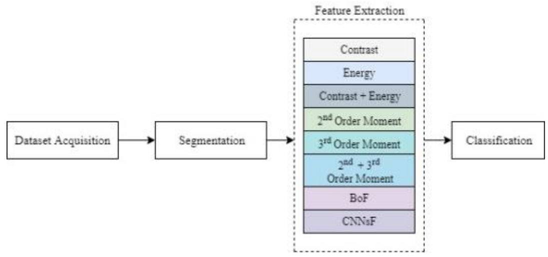


Figure 2: The block diagram of the experimental setup.

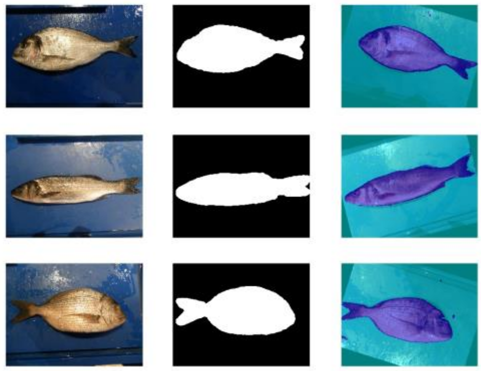
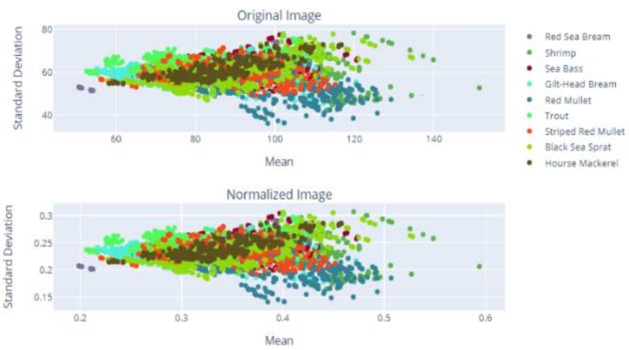


Figure 3 displays a series of images that present fish samples in a left-to-right sequence. These images are accompanied by their corresponding ground-truth segmentation masks and the segmentation outcomes produced by SegNet. The process of creating the ground-truth masks involves manual intervention from a human operator. Subsequently, morphological operators are applied to refine these masks. This refinement includes an erosion operation with a diamond-shaped structure and a distance of 8, followed by a dilation operation with a sphere shape and a distance of 25. These operations are uniformly applied to all masks, resulting in the final ground-truth masks. Once the ground-truth masks are obtained manually, the SegNet semantic segmentation algorithm [16] is used to automatically separate the fish from the background. SegNet is a neural network consisting of ten layers and is specifically employed for this purpose. The dataset is divided into a training set (70%) and a test set (30%). For the two classes, fish and background, a filter size of 3×3 and 64 filters are determined. The maximum epoch number is set to 10, and a mini-batch size of 8 is chosen. Due to the significant imbalance in pixel distribution between the fish and background, class weights are calculated using inverse frequency weighting. The last layer of the network is adjusted accordingly. Ultimately, SegNet achieves an average training accuracy rate of 98.01% and a test accuracy rate of 88.69%. Figure 3 provides visual examples of the ground-truth masks and SegNet results on the augmented dataset. Additionally, Table I presents detailed segmentation rates based on the Jaccard similarity index, expressed as percentages.).

After automatic segmentation, all seafood types were classified via the extracted features from segmented images. The information provided to the classifier are obtained through four different feature extraction methods as follows.

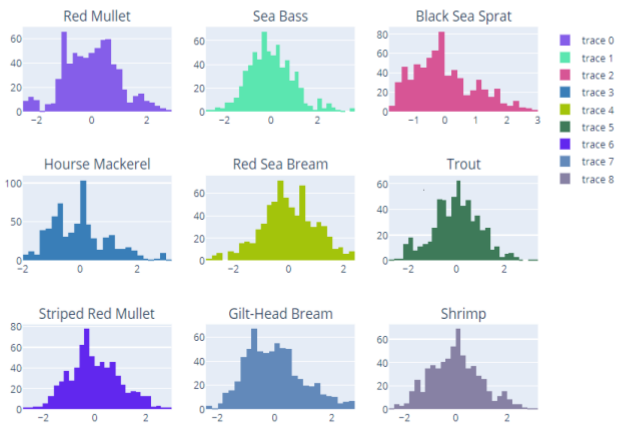
1) Mean: The mean, or average, of a set of numbers is obtained by summing all values and dividing by the total count.

2) Standard Deviation: The standard deviation measures the amount of variation in a dataset.



3) Z Test: When comparing two samples or a sample to a known population, a Z-test is a statistical technique used to see if there is a significant difference. Explore a particular hypothesis or compare metrics, you could apply a Z-test to your dataset on interactions with digital ads. Suppose you wish to run a Z-test on the daily time spent variable. Here is an example scenario to consider.

zscore=(x−Mean)/Standard Deviation



4) Logistic Regression : Logistic Regression is a statistical method for modelling the probability of a binary outcome. Unlike linear regression, which predicts a continuous outcome, logistic regression predicts the probability that an instance belongs to a particular category.

5) SVM: Support Vector Machine (SVM) classifier for a machine learning task using features extracted from a pre-trained neural network is used. Initially, an intermediate layer model is established to obtain feature vectors from the penultimate layer of an existing neural network ('model') for the training,

a) GLCM: GLCM presents how frequent distinct grayscale intensities of pixels appear in an image [17]. In this study, "contrast" and "energy" features are extracted through GLCM. While the contrast feature exploits the contrast between each pixel and its neighbour, the energy feature computes the sum of squared elements. The contrast feature can capture color variations of different fish species; hence the effect of distinct colors is aimed to be observed in the classification process. The energy feature on the other hand is chosen because of its proven efficiency in food quality assessment tasks [18]. These two features are fed to SVMs not only separately but also as a one concatenated feature.

b) BoF: The BoF algorithm is an adoption of the Bag of Words for image processing [19]. This feature extraction technique uses the SURF method to extract discrete pixel-based features to create a visual vocabulary [20]. After extracting all features, the weak ones are eliminated by using K-means clustering to obtain the final version of the visual vocabulary. Since it takes a considerable number of features into account, BoF is employed in this study.

c) CNNsF: CNNs are inspired by the human visual system and the biological structure of the brain and are specifically designed for image processing [18]. CNNs consist of four main layers; an input layer, a convolutional layer, a pooling layer, and a fully connected layer [21]. The number and order of layers are selected according to the difficulty of the problem at hand. More complex problems usually require a higher number of layers [22]. In this study, CNNsF is adopted to extract various features such as edges, blobs, and minute details [23]. Since training an end-to-end network requires high computational time, AlexNet is employed as a pre-trained network. The epoch number is selected as 10 and the minibatch size is determined as 8.

	Train Accuracy	Test Accuracy
Gilt Head Bream	97.12	96.82
Red Sea Bream	98.21	91.84
Sea Bass	95.69	82.98
Red Mullet	99.02	90.32
Horse Mackerel	98.90	86.97
Black Sea Sprat	97.52	89.66
Striped Red Mullet	99.24	89.60
Trout	97.32	80.45
Shrimp	99.06	89.59
Average	98.01	88.69

Table I: SegNet segmentation results in percentages.

C. Classification Results

Since first introduced in 1979 as a supervised binary classifier, SVMs [24] have been widely used in classification tasks. They are preferred as classifiers in this study because of their success in food assessment studies, e.g., [18], [22], [25]. In this set of experiments, one-versus-all SVM classifiers are designed for analysing the use of the extracted features. While randomly 70% of the images are used for training, 30% are employed to evaluate the algorithm. 10-fold cross-validation is used along with the SVMs and all feature vectors are normalized and formed with the same dimensions. Additionally, all experiments have been repeated five times with random initialization, and average values are provided in the experimental results. The convergence of computer vision and machine learning has had a tremendous influence on the food business during the last two decades. Industrial firms are aware of this impact. Automated techniques for assessing food quality have simplified procedures and improved the accuracy of spoiling detection. The seafood industry, which is critical for detecting degradation and infections, does not have publicly available statistics for quality analysis in supermarket retail areas. Fill this need, our study delivers a dataset of nine regularly eaten fish kinds in Turkey's Aegean Region. It was created with realistic restrictions in mind, and it features noisy backgrounds and varied exposures. Manual ground truth labelling, as well as semantic segmentation and feature-based classification tests, produce encouraging results. The dataset is now open to the public for further study. Notably, using contrast and energy features separately resulted in higher average accuracy than combining them.

Among statistical characteristics, the second-order moment consistently resulted in higher average accuracy rates. In training for horse mackerel and trout, it outperformed the GLCM contrast feature. In contrast to using these features individually, combining second and third-order moments resulted in a lower average accuracy.

The table shows the accuracy rates for various features in various seafood categories. Notably, CNNsF in SVMs outperformed BoF in terms of accuracy, closely matching the success of GLCM contrast and energy features in training. CNNsF produced the highest training and test accuracy rates for shrimp.

Table II: Feature based result for training of SVMs in percent.

	Contrast	Energy	BoF	CNNsF
Gilt Head Bream	98.00	97.26	81.67	93.07
Red Sea Bream	97.41	98.67	64.58	96.17
Sea Bass	95.56	96.37	75.00	92.08
Red Mullet	98.89	98.44	87.25	89.90
Horse Mackerel	97.41	98.15	80.33	92.74
Black Sea Sprat	97.85	96.41	93.33	96.07
Striped Red Mullet	97.33	97.85	65.00	86.66
Trout	99.19	97.15	95.83	95.00
Shrimp	97.15	96.96	90.95	97.56
Average	97.64	97.47	81.55	93.25

Table III: Feature based result for testing of SVMs in percent.

	Contrast	Energy	BoF	CNNsF
Gilt Head Bream	99.56	98.44	86.06	96.31
Red Sea Bream	98.85	99.33	66.61	97.86
Sea Bass	97.30	98.07	78.75	97.31
Red Mullet	99.56	99.19	92.14	97.62
Horse Mackerel	98.81	99.22	83.04	98.50
Black Sea Sprat	99.00	97.81	95.18	98.45
S Red Mullet	98.48	98.89	67.11	95.71
Trout	99.48	97.85	97.39	98.75
Shrimp	97.63	98.22	92.93	98.74
Average	98.74	98.56	84.36	97.81

III.CONCLUSION

The convergence of computer vision and machine learning has had a tremendous influence on the food business during the last two decades. Industrial firms are aware of this impact. Automated techniques for assessing food quality have simplified procedures and improved the accuracy of spoiling detection. The seafood industry, which is critical for detecting degradation and infections, does not have publicly available statistics for quality analysis in supermarket retail areas.

Fill this need, our study delivers a dataset of nine regularly eaten fish kinds in Turkey's Aegean Region. It was created with realistic restrictions in mind, and it features noisy backgrounds and varied exposures. Manual ground truth labelling, as well as semantic segmentation and feature-based classification tests, produce encouraging results. The dataset is now open to the public for further study.

REFERENCES

[1] FAO, "Global per capita fish consumption rises above 20 kilograms a year," 2016.

[2] J. Hu, D. Li, Q. Duan, Y. Han, G. Chen, and X. Si, "Fish species classification by color, texture and multi-class support vector machine using computer vision," *Comp. Electr. Agricult.*, vol. 88, pp. 133–140, 2012.

[3] R. Nijhawan, "A hybrid deep learning approach for automatic fish classification," *Proc. Emerging Trends Inf. Tech.*, p. 427, 2019.

[4] S. Hasija, M. J. Buragohain, and S. Indu, "Fish species classification using graph embedding discriminant analysis," in *Int. Conf. Mach. Vis. Inf. Tech.*, 2017.

[5] C. Spampinato, D. Giordano, R. D. Salvo, Y.-H. J. Chen-Burger, R. B. Fisher, and G. Nadarajan, "Automatic fish classification for underwater species behavior understanding," in *Proc. ACM Int. W. Anal. Retrieve. Tracked Events Motion Imagery Streams*, 2010.

[6] C. Pornpanomchai, B. Lurstwut, P. Leerasakultham, and W. Kitiyanan, "Shape and texture-based fish image recognition system," *Kasetsart J.-Natural Sci.*, vol. 47(4), pp. 624–634, 2013.

[7] S. O. Ogunlana, O. Olabode, S. A. A. Oluwadare, and G. B. Iwasokun, "Fish classification using support vector machine," *African J. Comput. & ICT*, vol. 8(2), pp. 75–82, 2015.

[8] R. Larsen, H. Olafsdottir, and B. K. Ersboll, "Shape and texture based. classification of fish species," in *Scandinavian Conf. Image Anal.*, 2009.

[9] S. Bermejo, "Fish age classification based on length, weight, sex and otolith morphological features," *Fisheries Res.*, vol. 84(2), pp. 270–274, 2007.

[10] S. Z. H. Shah, H. T. Rauf, M. Ikram Ullah, M. S. Khalid, M. Farooq, M. Fatima, and S. A. C. Bukhari, "Fish-Pak: fish species dataset from

Pakistan for visual features-based classification," *Data Brief*, vol.104565, 2019.
[11] B. V. Deep and R. Dash, "Underwater fish species recognition using. deep learning techniques," in *IEEE Int. Conf. Signal Process. Integrated Newt.*, 2019.
[12] B. J. Boom, P. X. Huang, J. He, and R. B. Fisher, "Supporting ground. truth annotation of image datasets using clustering," in *Proc. Int. Conf. Patt. Reclog.*, 2012.
[13] B. J. Boom, J. He, S. Palazzo, P. X. Huang, C. Beyan, H.-M. Chou, F.-P. Lin, C. Spampinato, and R. B. Fisher, "A research tool for long term and continuous analysis of fish assemblage in coral-reefs using. underwater camera footage," *Ecol. Inform.*, vol. 23, pp. 83–97, 2014.
[14] G. Cutter, K. Stierhoff, and J. Zeng, "Automated detection of rockfish in unconstrained underwater videos using haar cascades and a new image dataset: labeled fishes in the wild," in *IEEE Winter Appli. Comp. Vis. W.*, 2015.
[15] B. J. Boom *et al.*, "Long-term underwater camera surveillance for monitoring and analysis of fish populations," *Proc. Vis. Observation Anal. Animal Insect Behav.*, 2012.
[16] (Accessed May 6, 2020) SegNet. https://uk.mathworks.com/help/vision/ ug/semantic-segmentation-examples.html.
[17] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cyber.*, vol. SMC-3(6), pp. 610–621, 1973.
[18] O. Ulucan, D. Karakaya, and M. Turkan, "Meat quality assessment based on deep learning," in *Innov. Intell. Syst. Appli. Conf.*, 2019.
[19] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *W. Stat. Learn. Comp. Vis.*, 2004.
[20] E. Oyallon and J. Rabin, "An analysis of the SURF method," *Image Process. Online*, vol. 5, pp. 176–218, 2015.
[21] D. Karakaya, O. Ulucan, and M. Turkan, "Electronic nose and its applications: a survey," *Int. J. Automat. Comput.*, vol. 17, pp. 179–209, 2020.
[22] D. Karakaya, O. Ulucan, and M. Turkan, "A comparative analysis on fruit freshness classification," in *Innov. Intell. Syst. Appli. Conf.*, 2019.
[23] (Accessed May 6, 2020) Model. https://ww2.mathworks.cn/help/vision/ examples/image-category-classification-using-deep-learning.html.
[24] D. Boswell, "Introduction to support vector machines," *Dept. Comp. Sci. Eng., Univ. Calif. San Diego*, 2002.