

Classification of Shark Behavior from Accelerometry Data using Machine Learning

Sainesh Karan

*Dept. of Electrical Engineering
California State University, Long Beach*
Long Beach, USA

Emily Messe

*Dept. of Biological Sciences
California State University, Long Beach*
Long Beach, USA

Dr. Yu Yang

*Dept. of Chemical Engineering
California State University, Long Beach*
Long Beach, USA

Dr. Hen Guel Yeh

*Dept. of Electrical Engineering
California State University, Long Beach*
Long Beach, USA

Dr. Chris Lowe

*Dept. of Biological Sciences
California State University, Long Beach*
Long Beach, USA

Dr. Wenlu Zhang

*Dept. of Computer Science
California State University, Long Beach*
Long Beach, USA

Abstract—In this paper, we aim to ascertain the behavior of California Horn sharks (*Heterodontus francisci*) based on accelerometer data using a machine learning algorithm called Nearest Neighbors. We use digital signal processing techniques such as Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT) to represent the accelerometer signal in the frequency domain which helps to reduce the data size required for training the classifier, the computation time and the memory resources needed, besides improving model accuracy. The shark behavior is classified into four classes namely resting, swimming, feeding and non-directed motion. We compare different combinations of time and frequency domain data on the performance of the algorithm. It is shown that the transform domain data considerably improved the accuracy of the classifier.

Index Terms—Activity Recognition, Shark Behavior, KNN, Digital Signal Processing, Machine Learning

I. INTRODUCTION

Activity recognition pertains to the task of classifying physical activities undertaken by an individual from the collected data. Differentiating behaviors of free-ranging animals provides an insight into their feeding patterns that help us understand their ecology and their impact on the environment. Recently, accelerometers have been employed to unveil the abstruse lives of marine animals. It is extremely arduous to observe such species in their natural environments over long periods due to poor visibility, deeper depths, and adverse environmental conditions [1]. Over the years Bio-telemetry devices have become popular amongst researchers to overcome these obstacles and are widely available. An Acceleration Data Logger (ADL) is one such device that measures acceleration caused by body movements that can be used to determine body orientation and kinematics for behavioral classification [1]. ADLs allow data to be recorded and stored at high frequencies, and thus providing high-resolution data with useful information to reveal animal's behavior. However, the increased volume of high-resolution data is non-trivial to process, and the useful information hidden behind may not be easy to extract. This challenge motivates the research of automatic behavioral classification by using machine learning techniques.

The past few years have seen tremendous development in the fields of machine learning and artificial intelligence, which has spurred a widespread adoption of the technology for several applications. Machine learning algorithms can be broadly categorized into supervised and unsupervised methods. In supervised methods, a training data-set, which consists of the data and the associated outcome/labels, is used to train a model to accurately map the input variables to the outcome/labels. Once the model has been built based on the training data, it can be used to make predictions on new input data. Examples of such algorithms include Decision Trees, Random Forest (RF), K-Nearest Neighbors (KNN) and Neural Networks [1]. Several studies have been conducted in the past to study the behavior of free-ranging animals using data loggers and machine learning techniques. In [2] and [3], the authors demonstrate the utility of using data loggers and machine learning algorithms to classify the behavior of Griffon vultures and Cheetahs respectively. In [4] authors used machine learning techniques to identify the hunting and feeding pattern of penguins. Apart from behavior classification, authors of [5] and [6] have used data loggers to determine the mortality and energy expenditure of sharks respectively. In this paper, we apply the KNN algorithm to classify data collected from the ADL into four behavior classes namely Resting, Swimming, Feeding, and Non-directed motion. KNN is a simple machine learning algorithm in which, a new data point is classified based on the majority of its k - nearest neighbors, k being the user-defined parameter. The algorithm does not use any model to fit the new data sample but rather computes the distance of the data point from its neighbors [7]. We also explore the effect of signal transformation techniques such as Discrete Fourier Transform (DFT) and Discrete Cosine Transform (DCT) on the performance of the classification algorithm. We expect that the transformation techniques will enhance the differentiating features of each behavior class thereby aiding the algorithm in the process while reducing the computation complexity and time. All data pre-processing, simulations and machine learning techniques used in this paper were implemented in

II. DATA ACQUISITION AND PRE-PROCESSING

Laboratory trials were conducted to quantify acceleration signatures of California horn sharks (*Heterodontus francisci*) for different behaviors including resting, swimming, and feeding (prey capture and handling). Horn sharks were chosen as a model species to represent a demersal, resting shark that are capable of eating hard-shelled invertebrates [8]–[10]. Sharks were collected from Santa Catalina Island, and transported back to the CSULB Shark Lab. Trials occurred with a single shark in a 340 l tank with two GoPro video cameras, one mounted on the bottom of the tank, and one above the tank. For each trial, a shark was fitted with an accelerometer data logger (ADL; Cefas G6a+ or Technosmart AxyDepth) via a custom tag package. Accelerometers recorded 3D acceleration data at 25 Hz. All acceleration data were processed in IgorPro (vers. 6.2, WaveMetrics) using Ethographer [11]. Static acceleration represents the 3D body position and orientation due to gravity (i.e., pitch, X static; roll, Y static; yaw, Z static), whereas dynamic acceleration represents the 3D body movements (i.e., surge, X dynamic; heave, Y dynamic; sway, Z dynamic) as shown in Figure 1. Static acceleration was separated from dynamic acceleration using a 2 s box smoother [12]. Overall dynamic body acceleration (ODBA) was calculated by summing the absolute value of the dynamic axes and is used as a proxy for energetic expenditure and oxygen consumption [13]. Sharks were tagged approximately 6 hrs before the start of the trial to allow the shark to acclimate. To characterize motion patterns associated with prey manipulation and handling, instrumented sharks were then fed three different prey types including a live purple urchin (4 cm test diameter), a frozen and thawed squid (whole, with pen), and a frozen and thawed shrimp. Prey was buried or placed away from the shark in order to characterize natural feeding behaviors expected in the field. A trial concluded when the shark had consumed all trial prey. Trials were approximately 2 hrs long, and a total of 7 trials were done on 4 different individuals.

The data generated by the ADL was labelled manually by viewing the GoPro video recording of each trial and identifying the shark behaviour with the corresponding time-stamps. All data samples that were unlabelled were removed from the dataset. An example of the recorded data from dataset number 7 is shown in Fig.2 and Fig.3.

A. Data Pre-processing

A total of 7 datasets were formed from the trials. Each dataset was then separated into its four constituent behaviour classes based on the labels. The distribution of different classes in each dataset is presented in table 1. Since the datasets had very few feeding samples in comparison to other classes, careful consideration had to be given to how to split the data between training and testing. We chose dataset number 5 as testing data to test our model. This was done to maximise the number of feeding samples in the training data

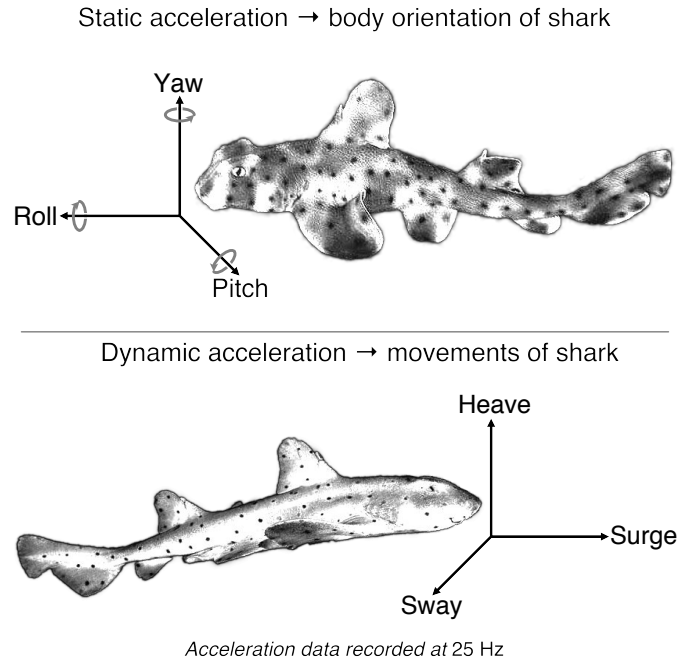


Fig. 1. Breakdown of acceleration data recorded at 25 Hz. Static acceleration represents the body position and orientation of the shark due to gravity. Dynamic acceleration represents the body movements of the shark.

while still having enough feeding samples to test the model. Also the dataset number 5 came from a different individual shark as compared to other datasets and would allow us to test capability of the model to generalise.

Next we combined all the datasets and subtracted the mean from all the data samples in the combined dataset so as to remove the DC-bias and noise, smoothing the overall dataset and centring the signal around zero [14]. We then normalized the all the datasets so as to standardize their mean and standard deviation to 0 and 1 respectively. As can be seen from Table 1, the trials were of varying lengths with varying distribution of classes. It can particularly be observed that the datasets are biased with the total feeding data samples being an order of magnitude lower than other classes. In order to overcome this bias, a subset of data samples was chosen from each of the remaining classes by examining the signals and choosing a subset that was most consistent and free of noise. A total of 100,000 samples was chosen for Resting behaviour with almost equal samples taken from each dataset. Similarly, 90,000 samples were chosen for swimming behaviour and 83,900 samples were chosen for non-directed motion. The distribution of the chosen subsets is presented in Table 2.

In order to aid the algorithm to differentiate between different classes from the input signal, we convert the input signals into different domains of representation. The repetitive nature of signals can be easily captured by using frequency domain techniques. This repetition often correlates to the periodic nature of a specific activity such as the tail beat of the shark while swimming or movement of the jaws while feeding. Among the most popularly used transformation technique is the Fourier

TABLE I
DISTRIBUTION OF DIFFERENT BEHAVIOR CLASSES IN THE DATASETS

	1	2	3	4	5	6	7	Total
Resting	157,750	565,580	379,850	10,250	6,150	27,975	77,374	1,224,929
Swimming	6,200	209,550	5,100	81,525	7,975	19,750	61,475	391,575
Feeding	5,700	350	200	1,375	875	2,900	2,100	13,500
NDM	51,950	7,400	15,700	1,750	1,775	27,025	5,400	11,1000
Total	22,1600	782,880	40,0850	94,900	16,775	77,650	146,349	

TABLE II
CLASS DISTRIBUTION OF CHOSEN SUBSETS USED FOR TRAINING

	Dataset 1	Dataset 2	Dataset 3	Dataset 5	Dataset 6	Dataset 7	Total
Resting	20,000	20,000	20,000	10,000	15,000	15,000	100,000
Swimming	6,000	20,000	5,000	20,000	19,000	20,000	90,000
Feeding	5,700	350	200	1,375	2,900	2,100	12,625
NDM	30,000	7,000	15,000	1,500	25,000	5,400	83,900
Total	61,700	47,350	40,200	32,875	62,900	42,500	

transform which allows one to represent the most important characteristics of a time-based signal such as its average (or DC component) and dominant frequency components in the frequency domain. Fourier transforms decompose an input signal into its constituent sinusoids. In addition to the DFT, other frequency-based representations have been used such as Discrete Cosine Transform which in contrast, decomposes a given signal into its constituent cosine waves. DCT returns an ordered sequence of coefficients such that the most significant information is concentrated at the lower indices of the sequence. This means that higher DCT coefficients can be discarded without losing information, making DCT ideal for data compression [15]. We compute the frequency analysis for the time signal of a specific length or window using the discrete Fourier transform with the Fast Fourier Transform (FFT) algorithm. In this case we use a window of 25 data samples or 1 second so as to match the sampling frequency of the ADL and to avoid the transitioning of behaviour classes within the window. In order to counter the bias in the training dataset against feeding data. We oversampled the feeding data by sliding the window with a 50 overlap [16] so as increase the number of feeding samples in the training dataset. Due to the symmetric property of Fourier Transform, we use only the 50% of the data in the frequency domain. Similarly, we use only 50% of the data of the DCT and discard the data in the higher frequency bins. However, it should be noted that the data reduction take place in when we use the transform domain data by itself. In the event where we combine the time and frequency domain data, the complete FFT and DCT are used so as to keep the input matrix to the model consistent. In such cases, no data reduction takes place. The total data samples in the transform domain are shown in table 3.

Thus by considering only the half of the Fourier transform i.e. ignoring the mirror image and the high frequency components of the DCT, we can reduce the total size of the data to be classified by half.

B. K-NN Classifier

K-Nearest Neighbor is a supervised learning algorithm where the result of new input query is classified as the same class as its k-nearest neighbors. Simplicity and low computational complexity are its main advantages which is why it has been widely used in several studies [17] [18][19]. The goal of this algorithm is to classify a new input based on attributes of the training samples. The training examples are vectors in a multidimensional feature space, each with a class label [18]. The training phase of the algorithm consists only of storing the data and class labels. There are different metrics to determine the distance between the input query and training samples. Euclidean, City Block, Chebychev are some of the common and popular distance metrics. The accuracy of the K-NN algorithm can be severely impacted by the presence of noisy input data. The values of k completely depend upon the nature of the data. A higher value of k reduces the effect of noise but reduces the distinction between the boundaries of the classes [18]. In cases where the total number of classes are even, k is selected to be odd to prevent ties during classification.

$$y = \operatorname{argmin}_{y=1\dots K} \sum_{j=1}^K P(i|x)C(i|j) \quad P(i|x) = \frac{k_i}{k} \quad (1)$$

In our study, in order to determine the optimum value of k we ran multiple iterations of the program with different distance metrics and values of k . It was observed that as we increased the value of k the accuracy of Resting and Swimming behaviors also increased but that of feeding and non-directed motion decreased. This may be since swimming and Resting classes have the most training samples and as the value of k is increased, the locality of the input query is destroyed and the algorithm starts to look at samples that are not neighbors. We were able to get the best accuracy of feeding and non-directed motion at $k = 15$ with the city block distance metric without sacrificing the accuracy of resting and swimming classes too much.

Another parameter that is used to improve the performance is the Cost Matrix. It is a square matrix where $\text{cost}(i,j)$ is the cost of classifying a point into class j if its true class is i (i.e., the rows correspond to the true class and the columns correspond to the predicted class of a confusion matrix). Naturally all diagonal elements of the cost matrix are zero since there no cost for correct classification. In our study, we developed the cost matrix by trial and error method. In the next section we compare the effect of the various pre-processing techniques discussed in section II on the performance of the classifier with and without the cost matrix.

$$\text{cost}(i,j) = \begin{bmatrix} 0 & 9.5 & 2 & 6 \\ 2 & 0 & 3.5 & 4 \\ 1.25 & 5 & 0 & 1.5 \\ 2.25 & 3 & 1.5 & 0 \end{bmatrix}$$

III. RESULTS

The results for the experiment without using the cost matrix are shown in Table 4:

Overall, this approach exhibits a wide range of accuracy results. Starting with only Time domain data it is observed that the technique surprisingly had the best accuracy for feeding behavior even though its overall accuracy was the lowest. The overall accuracy increases when the transform domain data is used for classification. It increases from 59.05% to 66.89% when using only FFT and to 67.39% after combining FFT and DCT. It however reduces back to 59.95% when time domain data is combined with transform domain. It can also be observed that DCT helps to increase feeding accuracy. The feeding accuracy jumps from 1.4% to 2.9% when FFT is combined with DCT. Similarly, the feeding accuracy jumps from 6.6% to 8.1% when time domain data is combined with DCT. The reason why DCT aids the improvement in feeding accuracy needs to be further investigated. For most cases the accuracy for Resting behavior remained around 50-55% only increasing up to 81% when using a combination of FFT and DCT. The swimming accuracy remained consistent in most cases only dropping to 63% when using only DCT data. This was rather surprising as the maximum training data samples belonged to this class. It was difficult to come to a conclusion on how the accuracy of the NDM behavior was impacted. It can however be observed that NDM accuracy is higher in cases where time domain data is used as opposed to transform domain data. On the whole the performance of the classifier was rather dismal as it struggled to classify feeding data accurately. This challenge was overcome by using a cost matrix to improve the feeding accuracy. The results for the classifier with the cost matrix are shown in table 5.

From table 5, It can be seen that while the overall accuracy is lower slightly lower than the previous case, but using a cost matrix vastly improves the accuracy of the Feeding behavior. Here the FFT and DCT combination had the highest accuracy for Resting and Feeding as well as the overall accuracy at 59.69% even though it showed a low accuracy of 13.6% for non-directed motion. Similar to the previous table we can

observe that DCT aids in improving the Feeding accuracy. However, DCT when used by itself, gave the lowest overall accuracy of 53.74%.

IV. CONCLUSION AND FUTURE WORK

In this paper we applied the K-NN algorithm to classify shark behavior based on acceleration data collected by an ADL. We used various frequency transformation techniques such as the FFT and DCT to study their impact on the performance of the algorithm. We also compare the effect of different combination of data in time and frequency domain on the algorithm's performance. We can conclude that the transforming raw data into frequency domain significantly improves the performance of the algorithm and that DCT was instrumental in improving the accuracy of Feeding Behavior. Although the reason for DCT improving the Feeding accuracy needs to be further investigated. We also successfully show that the signal transformation techniques can be used to reduce the overall data required for training the algorithm which results in a reduction in computational cost, memory requirement and computational time. However, there are several drawbacks in this study. A more robust methodology is needed to optimize the parameters of the algorithm instead of the trial and error method used in this study. Tuning the cost matrix is a rather challenging task and becomes even more difficult as the number of classes increase. Thus, a method to obtain the optimum cost matrix needs to be researched.

Future work may include a comparison between various other machine learning algorithms such as Support Vector Machines (SVM), Neural Networks and Decision Trees. Also feature extraction can be used to extract statistical features from the raw as well as frequency domain data to help the algorithm differentiate between the classes. The selection of features and the optimum number required could be the topic of further research. In addition to this future work may also include the use of wavelet analysis to derive time-frequency features from the data. Finally end goal of this research would be to use the models developed to study the data collected from sharks in the field so as to unravel their feeding patterns of California Horn sharks and derive an insight into their ecology.

REFERENCES

- [1] L. R. Brewster, J. J. Dale, T. L. Guttridge, S. H. Gruber, A. C. Hansell, M. Elliott, I. G. Cowx, N. M. Whitney, and A. C. Gleiss, "Development and application of a machine learning algorithm for classification of elasmobranch behaviour from accelerometry data," *Marine Biology*, vol. 165, no. 4, Aug 2018.
- [2] R. Nathan, O. Spiegel, S. Fortmann-Roe, R. Harel, M. Wikelski, and W. M. Getz, "Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures," *Journal of Experimental Biology*, vol. 215, no. 6, p. 986-996, 2012.
- [3] S. Grnew%elder, F. Broekhuis, D. W. Macdonald, A. M. Wilson, J. W. McNutt, J. Shawe-Taylor, and S. Hailes, "Movement activity based classification of animal behaviour with an application to data from cheetah (*acinonyx jubatus*)," *PLoS ONE*, vol. 7, no. 11, 2012.
- [4] S. Chessa, A. Micheli, R. Pucci, J. Hunter, G. Carroll, and R. Harcourt, "A comparative analysis of svm and idnn for identifying penguin activities," *Applied Artificial Intelligence*, vol. 31, no. 5-6, p. 453-471, Mar 2017.

- [5] N. M. Whitney, C. F. White, A. C. Gleiss, G. D. Schwieterman, P. Anderson, R. E. Hueter, and G. B. Skomal, "A novel method for determining post-release mortality, behavior, and recovery period using acceleration data loggers," *Fisheries Research*, vol. 183, p. 210–221, 2016.
- [6] A. C. Gleiss, R. P. Wilson, and E. L. C. Shepard, "Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure," *Methods in Ecology and Evolution*, vol. 2, no. 1, p. 23–33, Feb 2010.
- [7] S. Kaghyan and H. Sarukhanyan, "Activity recognition using k-nearest neighbor algorithm on smartphone with tri-axial accelerometer," *International Journal of Informatics Models and Analysis (IJIMA), ITHEA International Scientific Society, Bulgaria*, vol. 1, pp. 146–156, 2012.
- [8] D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. P. Cardoso, "Preprocessing techniques for context recognition from accelerometer data," *Personal and Ubiquitous Computing*, vol. 14, no. 7, p. 645–662, 2010.
- [9] M. A. Edmonds, P. J. Motta, and R. E. Hueter, "Food capture kinematics of the suction feeding horn shark, *heterodontus francisci*," *Environmental Biology of Fishes*, vol. 62, no. 4, pp. 415–427, 2001.
- [10] W. Strong, "Behavioral ecology of horn sharks, *heterodontus francisci*, at santa catalina island, california, with emphasis on patterns of space utilization (sharks)." 1990.
- [11] K. Q. Sakamoto, K. Sato, M. Ishizuka, Y. Watanuki, A. Takahashi, F. Daunt, and S. Wanless, "Can ethograms be automatically generated using body acceleration data from free-ranging birds?" *PLoS ONE*, vol. 4, no. 4, 2009.
- [12] E. Shepard, R. Wilson, F. Quintana, A. G. Laich, N. Liebsch, D. Albareda, L. Halsey, A. Gleiss, D. Morgan, A. Myers, and et al., "Identification of animal movement patterns using tri-axial accelerometry," *Endangered Species Research*, vol. 10, p. 47–60, 2008.
- [13] A. P. Summers, R. A. Ketcham, and T. Rowe, "Structure and function of the horn shark (*heterodontus francisci*) cranium through ontogeny: Development of a hard prey specialist," *Journal of Morphology*, vol. 260, no. 1, p. 1–12, 2004.
- [14] S. Sani, N. Wiratunga, and S. Massie, "Learning deep features for knn-based human activity recognition." 2017.
- [15] S. Preece, J. Goulermas, L. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, p. 871–879, 2009.
- [16] S. Rosati, G. Balestra, and M. Knaflitz, "Comparison of different sets of features for human activity recognition by wearable sensors," *Sensors*, vol. 18, no. 12, p. 4189, 2018.
- [17] M. Janidarmian, A. R. Fekr, K. Radecka, and Z. Zilic, "A comprehensive analysis on wearable acceleration sensors in human activity recognition," *Sensors*, vol. 17, no. 3, p. 529, Jul 2017.
- [18] M. Janidarmian, A. R. Fekr, K. Radecka, Z. Zilic, and L. Ross, "Analysis of motion patterns for recognition of human activities," *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare - Transforming healthcare through innovations in mobile and wireless technologies*, 2015.