Hierarchal Decomposition for Unusual Fish Trajectory Detection

Coral reefs are one of the most important natural environments which should be monitored to understand the environmental effects caused by global warming, pollution and so forth. Investigating such environments needs long-term monitoring and automatic analysis, although the traditional way is manual processing which is very labor intensive and time consuming.

Analyzing fish behavior is useful to detect environmental changes as fish behavior reflects environmental conditions. This analysis can be made by extracting the change in behavior pattern of fish or by finding abnormal behaviors (Beyan and Fisher, 2013). For instance, by analyzing the behavior of fish hovering over coral, the health of coral can be determined.

There are many video surveillance systems to observe fish behavior. The most well known way to analyze fish behavior is using video recordings where the camera is capturing the fish trajectories from a fish tank or in an aquarium (Papadakis, Papadakis, Lamprianidou, Glaroulos, & Kentouri, 2012). Diving to observe underwater using photography, hand-held video devices and optical systems are techniques that have been used to investigate fish behavior in natural environments. Acoustic systems, echo-systems and sonar have been used as well (Graham, Jones, & Reid, 2004). Alternatively, casting nets in the ocean and net casting with acoustic sensors are also popular to observe fish and determine their abundance (Spampinato et al., 2012). However, methods such as diving and net casting are not very suitable as they cause unusual fish behavior by frightening the fish. Moreover, with those approaches it is hard to capture huge amount of data and to do long-term monitoring (Spampinato et al., 2012). In recent years, as digital video recording systems become cheaper, collecting data in natural underwater environments with a fixed camera set up which is continuously recording underwater videos has become possible (Boom et al., 2014). Such a system results in massive amounts of underwater video although automatically and accurately analyzing data is still a challenging problem. At this point, computer vision techniques and pattern recognition methods could play an important role in analyzing the fish behaviors using underwater videos.

In the computer vision area, behavior understanding studies can be classified into two categories:

- Activity recognition,
- Unusual behavior detection (Piciarelli, Micheloni, & Foresti, 2008).

When the number of possible behavior models in an uncontrolled and uncooperative real-world is considered, activity recognition is very challenging as the system needs a definition of each activity (Piciarelli et al., 2008). As fish are usually not goal-oriented and make erratic movements due to water currents, the complexity of the movements' increases and makes encoding the behaviors into activities very challenging. On the other hand, unusual behavior detection analysis has become popular in recent years. To detect unusual behaviors, the system does not need any prior knowledge about the behaviors. The unusual behaviors are generally defined as outliers or rare events and are detected in an unsupervised fashion (Anjum and Cavallaro, 2008; Jiang, Yuan, Tsaftaris, & Katsaggelous, 2010).

In this study, we present an unusual fish trajectory detection system that analyzes natural underwater environment videos. The detection and tracking of fish is out of the scope of this study and the fish trajectories are obtained using the tool described in (Boom et al., 2014). The method proposed here classifies the trajectories as normal and unusual. Normal fish trajectories are defined as the trajectories which contain frequently observed behaviors while unusual trajectories are defined as the behaviors that are rare or outliers. The proposed method is a hierarchical decomposition method which is based on clustered and labeled training data where the similarity of data is used to build a hierarchy. Different from the research that used a fixed hierarchy based on features or classes and the methods that used the same feature set for classification of any class, we present a novel hierarchical decomposition which uses different feature and data subsets at different levels of the hierarchy. This allows more specific features to be used once the data focus onto specific subclasses. At this point, to understand the proposed method properly, its differences between previously proposed hierarchical decomposition methods and hierarchal classifiers should be identified. Hierarchical decomposition methods are generally applied to divide a multi class problem in a hierarchical way to obtain binary classes (Silla and Freitas, 2010). On the other hand, hierarchal classifiers uses pre-defined hierarchy such as decision trees and the classes are organized using this tree or a graph. The most similar work to ours was proposed by (Silla and Freitas, 2011). In that study (Silla and Freitas, 2011), different feature sets at different levels of the hierarchy was also used. However, a fixed taxonomy was applied for classification which distinguishes this study from ours.

The main contributions of this work are: *i)* presenting a novel approach for unusual fish trajectory detection which builds a feature or class taxonomy independent hierarchy, *ii)* demonstrating significantly improved performance on unusual fish trajectory analysis in unconstrained underwater videos, *iii)* given that the majority of works on unusual trajectory detection are unsupervised, the proposed method is different as being supervised and using labeled and clustered training data.

In the rest of this chapter, we first present a literature review on fish behavior understanding and unusual trajectory detection methods (Section 2). In Section 3, the proposed method is introduced. In Section 4, the experiments, used datasets, results and the comparisons with the state of art methods are given. Finally, in Section 5, we conclude the chapter with possible future works.

Recent Works

The definition of unusual behavior is a bit ambiguous in the literature. Unusual, abnormal, rare, outlier, suspicious, subtle, interesting and anomaly are words that can be used interchangeably depending on the application while they all refer to the uncommon behaviors (Morris and Trivedi 2008; Xu, Tang, Liu, & Zhang, 2010; Varadarajan and Odobez, 2009; Dickinson and Hunter, 2008; Jiang et al., 2010]. In this chapter, we prefer using the word “unusual” although reviews given below might use different wordings. On the other hand, behavior refers to trajectories for all sections of this chapter.

In this Section, we review studies on fish behavior analysis and works on unusual trajectory detection.

Review of Fish Behavior Understanding

Fish behavior monitoring studies which utilize computer vision and machine learning methods are becoming popular not only in biology but also in artificial intelligence. Existing studies mostly focus on water quality monitoring and toxicity identification using the behavioral stress responses of fish (Thida, Eng, & Chew 2009; Nogita, Baba, Yahagi, Watanabe, & Mori 1988; Schalie, Shedd, Knechtges, & Widder, 2001; Papadakis et al. 2012; Serra-Toro, Montoliu, Traver, & Hurtado-Melgar, 2010; Chew, Eng, & Thida, 2009). For instance, Thida et al. (2011) used trajectory shape features with a signed-distance function. Incremental spectral clustering was used to group the fish trajectories. Trajectories obtained from clean water were used to determine the abnormal trajectories in toxic water. Similarly, recurrence plots were used to analyze the swimming pattern of fish in the presence of chemicals in the water (Serra-Torro et al., 2010). The fish trajectories were represented as no movement, up, left, right and left movement and a string representing each trajectory was obtained using those movements. Strings were compared with Levenshtein and Hamming distances and used to build the recurrence-plots to detect abnormal swimming pattern. On the other hand, studies which consider other stress factors such as stocking density (Papadakis et al., 2012; Mancera et al., 2008) also exist. For instance, Papadakis et al. (2012) proposed a system to observe the behavior variability of *Sparus aurata* before and after feeding time during the day, the time that fish spent in inspecting the net and the number of bites on the net surface. The results in this study (Papadakis et al., 2012) showed that there is a connection between fish behavior, stocking density, and net condition. Fish feeding is influenced by stocking density and by the social interactions of fish.

Differently, automatically monitoring abnormal behavior of fish to help the farm operator in aquaculture sea cages (Pinkiewicz, Purser, & Williams, 2011) also exists. In that study, fish are tracked by a Kalman filter. 30 random fish were selected to analyze the fish behaviors in terms of the average swimming speed and direction. Normal and abnormal behaviors were distinguished by thresholding the values of calculated features.

A recent problem in this area is automatic fish motion pattern analysis in underwater environments (Spampinato et al., 2010; Amer et al., 2011; Beyan and Fisher, 2012; Beyan and Fisher, 2013). For instance, Spampinato et al. (2011) proposed an Adaptive Gaussian Mixture Model with the Adaptive Mean Shift algorithm to track fish in underwater. Texture and shape based features were used to recognize fish species. Lastly, fish trajectories were sub-sampled using the Douglass-Peucker algorithm and clustered using I-kMeans. This study can be seen as a preliminary work since it did not include any evaluation of the trajectory analysis. However, it is still important as it uses underwater videos and shows the importance of fish behavior analysis in that field. On the other hand, Amer et al. (2011) classified the underwater videos of fish using fish motion patterns. Fish behavior is modeled in terms of fish swimming speed, direction, periodicity and escape response time. Three sea depths were used and six behavior patterns were defined to identify a new video in terms of sea depth based on the behavior pattern. The most similar works to ours are (Beyan and Fisher, 2012; Beyan and Fisher, 2013) especially due to the trajectory dataset which contains unconstrained underwater videos. The former approach (Beyan and Fisher, 2012) tries to filter out normal trajectories to leave a more balanced normal and unusual trajectory set. It was applied to 2486 trajectories (to the best of our knowledge the second largest labeled fish trajectory dataset after the fish trajectory dataset used in this paper) which belong to 10 different fish species in Taiwanese Coral Reef. The results showed that the normal trajectory filtering rate of the method (Beyan and Fisher, 2012) is significant especially considering the behavior variations due to the different camera views and different fish species that were used in that paper. The later study (Beyan and Fisher, 2013) is a preliminary work of the proposed method in this chapter. It used a flat classifier with a single feature subset for classification of all trajectories therefore it is not based on the built hierarchy.

Some studies focused on the behavior of individual fish such as (Nogita et al., 1988; Schalie et al., 2001) when others studies considered fish schools (Thida et al., 2009; Chew et al., 2009). Some studies analyzed only one species like (Pinkiewicz et al., 2011; Chew et al., 2009; Kato et al., 2004; Xu, Liu, Cui, & Mioa, 2006). The majority of works analyzed the fish trajectories in a fish tank (Chew et al., 2009), aquarium (Thida et al., 2009) or an aquaculture sea cage (Pinkiewicz et al., 2011) which actually makes the analysis simpler as it decreases the number of fish behaviors, the variety of fish behaviors and most importantly eliminates the effects of habitat on the behavior of fish. A few studies worked on natural habitat underwater environments videos such as (Spampinato et al., 2010; Amer et al., 2011; Beyan and Fisher, 2012; Beyan and Fisher, 2013).

Reviews of Unusual Trajectory Detection Methods

Trajectories describe the displacements of objects and are typically considered as positions in 2 dimensions over time. Unusual trajectory detection studies can be categorized based on: *i)* the trajectory representation methods that they used (extracting multiple features such as velocity, acceleration, shape based features etc, using raw trajectory positions, or processed trajectory positions such as by polynomial fitting, Discrete Fourier Transform etc.), *ii)* the learning method that they used (unsupervised, supervised, semi-supervised).

Makris and Ellis (2002) used probabilistic Spline fitting to represent the trajectories, which was used to extract common pathways from a set of pedestrians' trajectories. Spline fitting does not need machine learning methods but the accuracy depends on choosing the correct number of control points. Brand and Kettnaker (2000) classified movement regions using an HMM based trajectory representation. The HMM is successful if the trajectory length is fixed for all trajectories given that each object detection represents a state in HMM. However, usually the lengths of trajectories are not equal. Therefore, to use HMM trajectory interpolation might be needed. Besides HMM based representation needs training data to define the states and transition matrixes. Principal Component Analysis (PCA) to represent segmented trajectories is used by Bashir et al. (Bashir, Qu, Khokhar & Schonfeld, 2005). In this work, trajectories were segmented into atomic actions using velocity and acceleration. PCA is useful as it provides a compact representation using eigenvectors but the number of components should be determined carefully as it is possible to lose a part of trajectory information. Sillito and Fisher (2008) used a fixed arc-length vector representation. The trajectory representation techniques: Haar wavelet coefficients, Discrete Fourier Transform (DFT), Chebyshev polynomial coefficients and Cubic B-spline control points were compared. These techniques were evaluated in terms of class separability since this metric is useful to evaluate an unusual trajectory detection method. Haar representation was found to be better than DFT while the highest separability values were obtained by Chebyshev or Spline representations. For more information, interested reader can refer to the survey on trajectory representations and similarity metrics (Morris and Trivedi, 2008).

Rather than explicitly reproducing the trajectories, the trajectories can be represented by the multiple features obtained from trajectories. For example, Zhong et al. (Zhong, Shi, & Visontai, 2004) used color and texture histograms. Behavior patterns are classified as normal and unusual using co-occurrence of these features. Porikli and Haga (2004) proposed to use object based and frame based features together to detect abnormal behaviors. In that study, object based features includes the histogram of aspect ratio, orientation, speed, color size of the object, the HMM trajectory representation, duration, length, displacement and global direction of the trajectory etc. As frame based features histogram of orientations, location, speed, size of objects etc. were used.

Unusual trajectory detection algorithms are commonly based on clustering and determine unusual trajectories as the trajectory that is not similar (close) to any known clusters using a pre-defined distance threshold or the trajectories that are similar to clusters that have few trajectories. For example, Hu et al. (2006) presented a hierarchical trajectory clustering method to detect abnormal trajectories and make behavior predictions. Position, velocity and size of the object were used to describe trajectories. At the first level of the hierarchy, trajectories are clustered using spatial information. At the second level, clustered trajectories are grouped according to temporal information. Abnormal trajectories were defined as the trajectories that belong to clusters having few samples. Self organizing maps (SOM) have also been used to detect unusual trajectories (Owens and Hunter, 2000). The trajectories were translated into a feature vector in terms of time smoothed positions and instantaneous velocity. The Euclidean distance between trajectories and clusters and a pre-defined distance threshold were used to find the unusual trajectories. A trajectory having a distance larger than a threshold becomes unusual. Another unsupervised unusual trajectory detection method was proposed by Izo and Grimson (2007). The normal and unusual trajectories were individually clustered using the Normalized Cuts Spectral Clustering algorithm. To represent the trajectories, a feature vector composed of the area of the object's bounding box, the speed, the direction of motion and the object position in the image were used. To classify a new trajectory, it is projected into the spectral embedding space of the obtained clusters and matched with the clusters. A 3-stage unsupervised hierarchical trajectory and activity learning process with an abnormal trajectory detection method was presented in (Morris and Trivedi, 2011). The trajectory points and the velocity extracted from the trajectory were used. In the first stage, interesting nodes were learned by GMM. In the second stage, the routes which represent each trajectory cluster were extracted using Longest Common Subsequence (LCSS) distance and spectral clustering. Following this, dynamics of activities were encoded using HMM. The abnormal trajectories were determined by comparing the trajectory's log-likelihood with a threshold.

In contrast to the studies using unsupervised methods, there are other unusual trajectory detection methods that utilize semi-supervised or supervised methods such as Support Vector

Machines (SVM) (Ivanov, Dufaux, Ha, & Ebrahimi, 2009), Hidden Markov Models (HMM) (Zhang, Gartica-Prez, Bengio, & McCowan, 2005), and Dynamic Bayesian Network (DBN) (Xiang and Gong, 2006; Loy, Xiang, & Gong, 2011). In these works, the methods use trajectories either fully labeled as normal and unusual (supervised methods) or only containing labeled normal trajectories (semi-supervised methods). For instance, velocity and acceleration features extracted from trajectories to detect unusual activities such as running or careless driving were used (Ivanov et al., 2009). In that study, SVM was applied and during training a model was learned using typical normal and unusual trajectories. The learned model was used to detect new unusual activities. As a different study, Xiang and Gong (Xiang and Gong, 2005) tried to find natural grouping of trajectories by eigenvectors of the behaviors' affinity matrix. Besides, they presented a time accumulative reliability measure to detect abnormalities. When the sufficient number of trajectories that belong to same behavior class is observed which is determined by the reliability measure, the normal trajectories were determined on-the-fly without manual labeling in order to detect the abnormalities. Behavior patterns were used to find the natural groupings and each group was represented by a DBN with Multi-Observation Hidden Markov Model (MOHMM) topology. For each new trajectory, the log-likelihood of it was determined by the MOHMM model. Then, all log-likelihoods were used to determine the abnormality of the trajectory using the reliability measure which is based on a threshold.

Proposed Method

The proposed hierarchy decomposition method utilizes *i)* clustering, *ii)* outlier detection and *iii)* feature selection to build the hierarchy. To automatically construct the hierarchy during training, clustering and outlier detection is combined with feature selection. The data is partitioned by using the selected features which are determined by feature selection, outlier detection and the ground-truth labels of the training data. In other words, the clustered and labeled data is used to determine the best feature set for the subset of training data in a certain level of the hierarchy. The details of the proposed method are given below.

Clustering

To partition the data we used Affinity Propagation (AP) (Frey and Dueck, 2007). Various studies have applied AP for clustering including anomaly detection. Unlike traditional clustering methods, AP determines cluster centers from the actual data samples which are called *cluster exemplars.* The method is based on the pair-wise similarity of the data samples where the negative of the Euclidean distance between data samples is used to define the similarity. There are two objective functions which include similarity calculations. One of them determines how appropriate it would be for data sample (*i*) to be the exemplar of another data point (*j*). The second one determines how appropriate it would be for the other data point (*j*) to choose data point (*i*) as its exemplar. The exemplars are the data points that maximize the overall sum of these two objective functions between all exemplars and exemplars' data samples. More information can be found in (Frey and Dueck, 2007).

There are many reasons to prefer AP over traditional clustering methods (k-means, hierarchical clustering etc.). The main reasons here are its ability to produce smaller clusters and the ability to produce uneven sized clusters which is compatible with the outlier detection method that we propose. Additionally, its fast processing speed makes the proposed method faster. Being non-parametric, not requiring initialization and not depending on sample order makes using a validation set unnecessary and helps to reduce training time. Its scalability also makes the proposed classification algorithm scalable as well.

Outlier Detection

An outlier is generally defined as data sample that is far from the other data samples in the same cluster. The numbers of outliers are smaller than the numbers of other data samples in the same cluster. In the context of the work presented here, unusual trajectories are what we want to discover and outlier detection is used to detect unusual trajectories.

Motivated by the study on trajectory clustering (Anjum and Cavallaro, 2008), two types of outliers are defined:

- Outliers located in small clusters,
- Outliers located in dense clusters but distant from the cluster exemplar

The small and dense clusters are identified using the cardinality of the clusters. A cluster which has fewer trajectories than 10% of the median cardinality of the clusters or a cluster that has only one trajectory is defined as a small cluster. All trajectories belong to such a cluster are classified as unusual trajectories. If the cluster is not a small cluster, then the unusual trajectories (outliers) are detected using the Euclidean distance between the trajectory and the cluster exemplar. If the calculated distance is further than the threshold $\tau=\mu+w\sigma$ (μ: mean, w: weight and σ: standard deviation of all distances between all trajectories and cluster exemplar) of that cluster, then that trajectory is classified as an outlier (unusual trajectory). Otherwise it is classified as a normal trajectory. As seen, this threshold is specific for each cluster as it is calculated in terms of the properties of the cluster such as mean, standard deviation of the distances between trajectories and the cluster exemplar. The w is chosen as it is given in Results section. On the other hand, evolutionary algorithms can be adapted to find the optimal *w* but in our experiments, the values of *w* that we used were good enough to obtain good performances.

Feature Selection

Feature selection is integrated with clustering and outlier detection. The advantage of feature selection is to prevent over-fitting, eliminate irrelevant and redundant features and the features which might misguide classification (Pudil, Novovicova, & Kittler, 1994). Sequential Feature Selection (Pudil et al., 1994) is used to determine the best feature sets at each level of the hierarchy. As the feature selection criterion the mean of the true positive rate (*TPrate*) and the true negative rate (*TNrate*) as defined in experiments and result part are used. The traditional feature selection criterion (accuracy: total number of correctly detected trajectories over total number of trajectories) was not applied as it increased the misclassification of the unusual trajectories.

Feature selection is applied as follows:

- Given an empty feature set, clustering and outlier detection are applied to the data using each feature individually. The mean of *TPrate* and *TNrate* is calculated using the ground-truth data. The set giving the highest mean of *TPrate* and *TNrate* determines the current feature set (which currently has single feature).
- Given the remaining set of features and the current feature set, an additional feature is added by applying the same procedure. After all possible additional features are tried; the extended feature set which gives the best performance is kept.
- Adding features to the current feature set continues until the classification performance decreases compared to the previous feature subset.

Hierarchy Decomposition

At each level of the hierarchy, using the best feature set found by feature selection, the data is clustered using AP. Outlier detection is applied to each cluster individually and the unusual trajectories at the current level of the hierarchy are found. Then, using the ground-truth data for each cluster, misclassified normal or unusual trajectories are found (if they exist). The clusters which do not contain any misclassified trajectory are kept for that level, and the corresponding trajectories are not used for construction of the rest of the hierarchy. Such clusters are called "perfectly classified clusters". On the other hand, clusters which have at least one misclassified trajectory no matter whether unusual or normal are used to continue the hierarchy construction. Using the clusters that have misclassified trajectories, the hierarchy construction recurses in the same way. By repeating clustering, outlier detection and feature selection, the hierarchy construction continues until there is no cluster which is perfectly classified or all trajectories are perfectly classified. In summary, at each level of the hierarchy, different trajectories are used and to distinguish those trajectories, different feature subsets are utilized. Once a trajectory that belongs to perfectly classified cluster at any level of the hierarchy is detected, it is never used for hierarchy construction in the next levels.

The leaf nodes of the hierarchy contain either: perfectly classified clusters (can mostly observed at the upper levels of the hierarchy) or misclassified clusters (can only be observed in the leaf nodes belong to the last level of the hierarchy).

A cluster called perfectly classified can be either:

- Perfectly classified mixed cluster: Contains unusual and normal trajectories. All trajectories are correctly classified using the outlier detection threshold.

- Perfectly classified pure normal cluster: A dense cluster which contains only normal trajectories which are correctly classified using the outlier detection threshold.
- Perfectly classified pure unusual cluster: Contains only unusual trajectories which are correctly classified since being in small clusters and we assume that small clusters contain only unusual trajectories.

A cluster called misclassified can be either:

- Misclassified mixed cluster: A dense or small cluster which contains both unusual and normal trajectories with at least one trajectory wrongly classified using the outlier detection threshold.
- Misclassified pure normal and dense cluster: Contains only normal trajectories with at least one trajectory wrongly classified as an unusual trajectory using the outlier detection threshold.
- Misclassified pure normal and small cluster: Contains only normal trajectories with at least one trajectory wrongly classified as unusual trajectory due to being in a small cluster.
- Misclassified pure unusual cluster: A dense cluster that contains unusual trajectories with at least one trajectory wrongly classified as normal trajectory using the outlier detection threshold.

The hierarchy construction algorithm is illustrated in Figure 1.

New Trajectory Classification Using the Constructed Hierarchy

A new trajectory is classified using the constructed hierarchy with all perfectly classified clusters and misclassified clusters at all levels, the selected feature subsets for each level and the outlier detection thresholds for each cluster. It is based on finding the closest clusters at each level of the hierarchy. The closest cluster is found using the Euclidean distance between the new trajectory and the cluster exemplars with the selected features for that specific level, including misclassified clusters as well. Therefore, at each level in the hierarchy, the closest cluster can be one of the cluster types given in hierarchy construction. Based on the closest cluster and the position in the closest cluster, the classification of the new trajectory can be: unusual trajectory, candidate normal trajectory or no effect on the decision, as given in Table 1.

In summary, we used the heuristic that even a single level's decision as unusual trajectory is enough to classify the new trajectory as an unusual trajectory no matter what the level of the hierarchy is. Any decision that the trajectory is candidate normal makes the new trajectory go to the next level to also be evaluated in there. If there is no decision as an unusual trajectory from any level and if the decision of at least one level is candidate normal then the final class of the new trajectory is determined as normal trajectory. A decision of "no decision" does not have any effect on the classification of the new trajectory. However, it is possible that the closest cluster at each level of the hierarchy is a misclassified cluster. In this case, the ground-truth labels of the training trajectories are used to apply the rules given in Table 2. The rules given in Table 1 and 2 are illustrated in Figure 2.

Other decision heuristics can be applied other than the heuristic that we use (decision as an unusual trajectory at any level stops classification of the new trajectory while a decision as a normal trajectory sends the new trajectory to the next level). For instance, the inverse heuristics: any decision as normal trajectory stops classification regardless of the level of the hierarchy while decision as an unusual trajectory send the new sample to the next hierarchy level can be applied. Alternatively, majority voting using the decisions from each level can determine the final class of the new trajectory.

Experiments and Results

The proposed method was compared with the state of art classification algorithms, outlier detection methods and trajectory analysis methods. The evaluations were performed using a fish trajectory dataset and a pedestrian dataset in terms of *TPrate* (unusual trajectory detection, Eq. 1), *TNrate (*normal trajectory detection Eq. 2) and geometric mean of *TPrate* and *TNrate* (GeoMean: represents overall detection, Eq. 3). The positive class represents unusual trajectories and negative class represents normal trajectories. The GeoMean is preferred as it does not ignore the importance of the classification of unusual trajectories due to being under-represented and also is a suggested metric for imbalanced datasets (Kubat and Matwin, 1997).

$$\text{True Positive Rate (TPrate)} = TP / (TP+FN) \tag{1}$$

$$\text{True Negative Rate (TNrate)} = TN / (TN+FP) \tag{2}$$

$$GeoMean = \sqrt{TPrate \ x \ TNrate} \tag{3}$$

In Eqs. 1-3; *TP* is the number of correctly classified unusual trajectories, *TN* is the number of correctly classified normal trajectories, *FN* is the number of misclassified unusual trajectories and *FP* is the number of misclassified normal trajectories.

Datasets

The fish trajectory dataset we used was the set presented in (Beyan and Fisher, 2013b) which includes 3120 fish trajectories all belong to *Dascyllus reticulatus* observed in the

Taiwanese Coral Reef (http://groups.inf.ed.ac.uk/f4k/GROUNDTRUTH/BEHAVIOR/). This set contains 3043 normal and 59 unusual trajectories with 179 features. This set is preferred as it is the largest public fish trajectory dataset and each trajectory has its class labels as well (Beyan and Fisher, 2013b). Examples of normal and unusual trajectories are given in Figure 3.

In addition, the proposed method was applied to a pedestrian trajectory dataset as well. The data which belongs to 1st of September (one of the largest set, having 1634 normal, 718 unusual trajectories) in the Forum Pedestrian database (Majecka, 2009) was utilized. As features, similar to fish trajectory dataset we extract the features: acceleration based, vicinity based, curvature scale space based, centre distance function in 2 dimension, loop, moment based, turn, and velocity based as presented in (Beyan, and Fisher, 2013). Additionally, trajectory points after B-spline fitting and the difference between the B-spline fitted trajectory and the real trajectory points were also used. Altogether 758 features are obtained. To prevent possible over-training or the curse of dimensionality, Principal Component Analysis (PCA) is applied to each group of features individually. To define the number of components for PCA, the smallest number of component that represents the 90% of the sum of all eigenvalues is used. As a result, 57 PCA features were obtained.

Results

The results presented in this section can be divided into 2 subsections: *i)* comparisons with the state of art methods and *ii)* evaluation of different heuristics for classification of new fish trajectory.

For all experiments presented in this section, 9-fold cross validation was performed. Training, validation and test sets were constituted randomly. The normal and unusual trajectories are distributed equally in each set. For the methods using sequential forward feature selection, validation sets are used to pick the best feature set for each method individually. For others including the proposed method (constructs the hierarchy only using the training set) validation sets were not used. The training and testing sets are kept the same for all methods.

Comparisons with the State of Art Methods

The proposed method is compared with the following methods with the given settings:

- **k- Nearest Neighbors (kNN):** *k* were used as *{1, 2, 3, 4, 5, 10, 15, 25}*.
- **k- Nearest Neighbors with Feature Selection (kNN-wFS):** The same *k* values with the kNN were used while sequential forward feature selection was applied.
- **Support Vector Machines (SVM):** Radial basis function with varying kernel parameters was used as the kernel function. Sequential Minimal Optimization was used to separate hyperplanes. All features are used for detecting unusual trajectories.
- **SVM with Feature Selection (SVM-wFS):** Applied as given in SVM description but integrating with sequential forward feature selection.
- **Random Forest with Balanced Training (RF-BT):** The trees are grown without pruning. A number of trees *{10, 30, 50, 70, 100, 120, 150, 200, 500, 1000}* were tested. For node splitting, the Gini index (Breiman, Friedman, Olshen, & Stone, 1984) was used. For balance training, all unusual trajectories were kept, and subsets of the normal trajectories were chosen randomly. Therefore, the numbers of normal trajectories in the chosen subset become equal to the numbers of total unusual trajectories. All features are used to detect unusual trajectories.
- **RF-BT with Feature Selection (RF-BT-wFS):** Applied as given in RF-BT description but integrating with sequential forward feature selection.

- **Unsupervised Modeling of Object Tracks (Izo and Grimson, 2007) (UMOT):** Normal and usual trajectories are clustered individually by normalized cuts spectral clustering. Each cluster was modeled as a mixture of Gaussians in the spectral embedding space. A new trajectory is classified using the likelihood by projecting it into the spectral embedding space from normal and unusual classes. Different sigma values such as *{1, 10, 20 etc.}* and different cluster sizes *{10, 15, 20, 30, 40, 50, 60, 80, 90}* for normal and usual clusters were tested.

- **Local Outlier Factor (Janssens, 2009) (LOF):** This method assumes that if there are not many samples in the surrounding space of a trajectory, then that trajectory is an outlier. Clustering is not needed. Training is performed only using normal classes. During validation normal and unusual class trajectories are used and the best feature set is selected using sequential forward feature selection. The neighborhood is defined with a parameter taken as *{1, 3, 5, 10, 15, 20 and 25}*.

- **Filtering method (Beyan and Fisher, 2012) (Filtering):** The search area pixel values were taken as *{2, 4, 8, 16, 20}*.

- **Flat Classifier (Beyan and Fisher 2013):** Outlier detection parameter w was taken as *{-1, -0.3, 0, 0.3, 0.6, 0.9, 1, 2, 3, 6}*.

- **Proposed Method (Proposed):** Outlier detection parameter w was taken as {0, 0.3 and 1} for the fish trajectory dataset and {-1, 0, 0.3, 0.6, 1 and 2} for Forum Pedestrian Database (Majecka, 2009).

In Table 3, the best results for GeoMean and the corresponding *TPrate, TNrate* using the fish trajectory dataset are given. For each evaluation metric the standard deviation (considering cross validation folds) is also given after ± sign. The best results of each evaluation metric are emphasized in bold-face.

The results show that the proposed method has highest unusual fish trajectory detection rate (*TPrate*) and is also the best method overall (*GeoMean*). For the proposed method the best performance was observed when the outlier detection threshold w is 0. The depth of the hierarchy was at most 3 while mostly 2 for the 9-folds. Paired *t-tests* were applied to the *GeoMean* data between each other method and the proposed method. It is found that the proposed method is significantly better than all methods except RF-BT, RF-BT-wFea and SVM ($\alpha=0.05$).

To show that the method is not limited to fish trajectory analysis but a general unusual trajectory detection method as well, we applied it to the pedestrian trajectory dataset (Majecka, 2009). The performance of the proposed method is compared with RF-BT, RF-BT-wFea, SVM-wFea as they performed well on the fish trajectory dataset. Also, LOF (Janssens, 2009) was compared since this method is one of the most popular outlier detection methods and was applied in (Hsiao, Xu, Calder, & Hero) as one of the state of art method for that dataset.

The best results for GeoMean and the corresponding *TPrate, TNrate* using pedestrian trajectories are given in Table 4. For this dataset the best performance of the proposed method was observed when the outlier detection threshold w is 0.3. The depth of the hierarchy was at most 5 while mostly 3 for the 9-folds. For this dataset, the proposed method performed the best to detect unusual trajectories (*TPrate*) and also in terms of *GeoMean*. A paired *t-test* applied between each method and the proposed method using the GeoMean results showed that the proposed method is significantly better than each other method ($\alpha=0.05$).

Evaluation of Different Heuristics for Classification of New Fish Trajectories

The proposed method is compared with variations of the algorithm using different heuristics to classify the new trajectories (Alter1-4). The contribution of having different levels with different subsets of trajectories and features is explored by applying the proposed method to all selected features from different levels as they are selected in a single level including all training trajectories (SingleLevProposed). Additionally, the features selected by the proposed method are evaluated by using SVM classifier (SVMwPropFea). The contribution of outlier detection algorithm is tested by keeping the same heuristics but changing the decision maker as SVM (Hie-SVM, Hie-SVM-Alter1). All those methods are defined below in detail and the best results in terms *GeoMean* with corresponding *TPrate*, *TNrate* are given in Table 5.

The different heuristics used to classify new fish trajectories are:

- **Proposed Method (Proposed):** Outlier detection parameter w was taken as *{0, 0.3 and 1}*.The heuristic is: a decision as a "**unusual trajectory**" at any level stops the classification of the new trajectory and the new trajectory become unusual, while a decision as a "**normal trajectory**" sends the new trajectory to the next hierarchy level.

- **Single level classification using features selected by Proposed (SingleLev Proposed):** The proposed method was applied using all the features selected (without feature selection) from all levels during hierarchy construction of proposed method. The classifier is outlier detection but the new hierarchy has only one level. Outlier detection parameter *w* was taken as *{0, 0.3 and 1}*.

- **Alternative Heuristic 1 (Alter1):** Outlier detection parameter w was taken as *{0, 0.3 and 1}*.The heuristic is: a decision as a "**normal trajectory**" at any level stops the classification of the new trajectory and it become normal, while a decision as a "**unusual trajectory**" sends the new trajectory to the next hierarchy level.

- **Alternative Heuristic 2 (Alter2):** Find the closest cluster at each level using corresponding features. Then, find the closest cluster of all which might be from any level of the hierarchy. If the closest cluster is a perfectly classified cluster then, a decision as unusual trajectory makes the new trajectory unusual and a decision as normal trajectory makes the new trajectory normal. If the closest cluster is a misclassified cluster then, the ground-truth labels are used as Proposed applies. The outlier detection parameter *w* was taken as *{0, 0.3 and 1}*.

- **Alternative Heuristic 3 (Alter3):** Apply the **proposed method**, classify the new trajectory at every level of the hierarchy and combine the decisions using **majority voting**. If the numbers of levels classifying the new trajectory as unusual and normal are equal, then the new trajectory is **unusual**. The outlier detection parameter *w* was taken as *{0, 0.3 and 1}*.

- **Alternative Heuristic 4 (Alter4):** Apply the **proposed method**, but classify the new trajectory at every level of the hierarchy and combine the decisions using **majority voting**. If the numbers of levels classifying trajectory as unusual and normal are equal then the new trajectory is **normal**. The outlier detection parameter *w* was taken as *{0, 0.3 and 1}*.

- **SVM using features selected by Proposed (SVM-wPropFea):** The features selected by Proposed in all levels are utilized in a single SVM classifier. SVM was applied with the settings given above.

- **Hierarchical SVM (Hie-SVM):** Applying **Proposed** but using SVM as the classifier instead of the outlier detection algorithm. SVM was applied with the settings given above.

• **Hierarchical SVM- Alternative Heuristic 1 (Hie-SVM-Alter1):** Applying **Alter1** but using SVM as the classifier instead of the outlier detection algorithm. SVM was applied with the settings given above.

As seen in Table 5, the proposed method is the best in terms of *GeoMean* and *TPrate*. SVMwPropM3Fea also performed well which means that the selected features by proposed method are representative to detect unusual fish trajectories. SingleLevProposed did not perform as well as the proposed method which means that utilizing different features for different trajectory subsets is more successful. Alter1 and Alter4 did not perform as well as proposed method and Alter2 and Alter3. That is because their *TNrate* were not as good as their *TPrate* which decreased the *GeoMean* as well. Hie-SVM did not perform significantly worse than the proposed method but on average the proposed method is better with higher *TPrate*. Similar to Alter 1, Hie-SVM- Alter1 also tended to classify samples as the normal class therefore its *TNrate* is greater than the *TNrate* of Hie-SVM but its *TPrate* is much worse which makes its *GeoMean* worse than the *GeoMean* of Hie-SVM.

Conclusions and Future Works

In this chapter, we presented a hierarchical decomposition method which constructs the hierarchy based on clustered and labelled trajectories using the similarity of trajectories. Different feature sets applied to different subset of trajectories formed the hierarchy.

The results showed that the proposed method had a better performance compared to the state of art classification methods and unusual trajectory detection methods and especially in terms of the unusual trajectory detection rate. Besides, its high normal trajectory detection rate is helpful for marine biologists since it allows filtering out many normal trajectories with a low error rate and lets them to focus more on unusual trajectories which are important given that they have huge amounts of data. The proposed algorithm's performance was also validated by another trajectory dataset. Moreover, the proposed method is also computationally efficient at classifying a new trajectory as it is only based on distance calculations while traversing the built hierarchy.

The proposed method can be applied for classification of binary imbalanced data sets including environmental data sets given that it is not limited to unusual fish trajectory detection. On the other hand, the proposed method can be considered closed to bagging since it does not use all the data samples to build up the hierarchy at each level. However, in our case, the bags are defined by the performance of the classifier (we continue to build up the hierarchy with the misclassified trajectories) but not as random subsets as happens in bagging. Moreover, it is different from boosting by using a subset of data in addition to not using a weight to support the classification of misclassified trajectories.

In the future, we will investigate the performance of the proposed method on imbalanced datasets from various application areas using the different heuristics that were presented. Moreover, the proposed method will be applied to larger fish datasets which might also include other fish species.

References

Amer, M., Bilgazyev, E., Todorovic, S., Shah, S., Kakadiaris, I., and Ciannelli, L. (2011). Fine-grained Categorization of Fish Motion Patterns in Underwater Videos, *In IEEE International Conference on Computer Vision Workshops*, 1488-1495.

Anjum, N., and Cavallaro, A. (2008). Multifeature Object Trajectory Clustering for Video Analysis, *IEEE Transactions on Circuits and Systems Video Technology*, 18(11), 1555-1564.

Bashir, F., Qu, W., Khokhar, A., and Schonfeld, D. (2005). HMM-based Motion Recognition System Using Segmented PCA, *In Proceedings of IEEE International Conference on Image Processing (ICIP),* 1288-1291.

Beyan, C., and Fisher, R. B. (2012). A filtering mechanism for normal fish trajectories. *In Proceedings of IEEE International Conference on Pattern Recognition (ICPR),* 2286–2289.

Beyan, C., Fisher, R. B. (2013). Detecting abnormal fish trajectories using clustered and labeled data, *In Proceedings of IEEE International Conference on Image Processing (ICIP),* 1476 - 1480.

Beyan, C., Fisher, R. B. (2013b). Detection of Abnormal Fish Trajectories Using a Clustering Based Hierarchical Classifier, *In Proceedings of British Machine Vision Conference (BMVC), 22.1-22.11*.

Boom, B. J., He, J., Palazzo, S., Huang, P. X., Beyan, C., Chou, H., Lin, F., Spampinato, C., and Fisher, R. B. (2013). Research tool for the analysis of underwater camera surveillance footage, *Ecological Informatics*, 2383-2397.

Breiman L., Friedman J.H., Olshen R.A., and Stone C.J. (1984). Classification and Regression Trees. Wadsworth and Brooks, Monterey, CA.

Chew, B. F., Eng, H. L., Thida, M. (2009). Vision-based Real-time Monitoring on the Behavior of Fish School, *In Proceedings of International Association for Pattern Recognition Conference on Machine Vision Applications*, 3(16), 90-93.

Dickinson, P. and Hunter, A. (2008). Using Inactivity to Detect Unusual Behaviour, *IEEE Workshop on Motion and Video Computing*, 1-6.

Frey, B. J., and Dueck, D. (2007) Clustering by Passing Messages Between Data Points, *Science*, 315, 972-976.

Graham, N., Jones, E. G., and Reid, D. G. (2004). Review of Technological Advances for the Study of Fish Behavior in Relation to Demersal Fishing Trawls, *ICES Journal of Marine Science*, 61, 1036-1043.

Hu, W., Xiao, X., Fu, Z., Xie, D., Tan, T., and Maybank, S. (2006). A System for Learning Statistical Motion Patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 1450- 1464.

Hsiao K., Xu K., Calder J., and Hero A. O. (2012). Multi-criteria Anomaly Detection Using Pareto Depth Analysis, *In Proceedings of Neural Information Processing Systems (NIPS)*.

Janssens J.H.M. (2009). Outlier detection with one-class classifiers from ML and KDD. *In Proceedings of International Conference on Machine Learning Applications,* pages 147–153.

Jiang, F., Yuan, J., Tsaftaris, S. A., and Katsaggelous, A. K. (2010). Video Anomaly Detection in Spatiotemporal Context, *In Proceedings of IEEE International Conference on Image Processing (ICIP)*, 705-708.

Ivanov, I., Dufaux, F., Ha, T., M., and Ebrahimi, T. (2009). Towards Generic Detection of Unusual Events in Video Surveillance, *In Proceedings of Advanced Video and Signal Based Surveillance (AVSS)*, 61-66.

Izo T. and Grimson W.E.L. (2007). Unsupervised modeling of object tracks for fast anomaly detection. *In Proceedings of IEEE International Conference on Image Processing (ICIP)*, 529–532.

Kato, S., Nakagawa, T., Ohkawa, M., Muramoto, K., Oyama, O., Watanabe, A., Nakashima, H., Nemoto, T., and Sugitani, K. (2004). A computer image processing system for quantification of zebrafish behavior. *Journal of Neuroscience Methods*, 134, 1–7.

Kubat, M., and Matwin, S., (1997). Addressing the curse of imbalanced training sets: one-sided selection, *In Proceedings of International Conference on Machine Learning (ICML),* 179–186.

Loy, C. C., Xiang, T., and Gong, S. (2011). Detecting and discriminating behavioural anomalies, *Pattern Recognition*, 44, 117- 132.

Majecka, B. (2009). Statistical Models of Pedestrian Behaviour in the Forum, *Master's Thesis, School of Informatics, University of Edinburgh*.

Makris, D., and Ellis, T. J. (2002). Spatial and Probabilistic Modelling of Pedestrian Behaviour, *In Proceedings of British Machine Vision Conference (BMVC)*, 2, 557-566.

Mancera, J. M., Vargas-Chacoff, L., Garcia-Lopez, A., Kleszczynska, A., Kalamarz, H., Martinez Rodriguez, G., Kulczykowska, E. (2008). High density and food deprivation affect arginine vasotocin, isotocin and melatonin in gilthead sea bream (Sparus auratus). *Comparative Biochemistry and Physiology Part A*, 149, 92–97.

Morris, B. T. and Trivedi, M., M. (2008). A Survey of Vision-based Trajectory Learning and Analysis for Surveillance, *IEEE Transactions on Circuits and Systems for Video Technology,* 18(8), 1114- 1127.

Morris B. T. and Trivedi. M. M. (2011). Trajectory learning for activity understanding: Unsupervised, multilevel and long-term adaptive approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11), 2287–2301.

Nogita S., Baba K., Yahagi H., Watanabe S., and Mori S. (1988). Acute toxicant warning system based on a fish movement analysis by use of AI concept. *In Proceedings of IEEE International Workshop on Artificial Intelligence for Industrial Applications*, 273–276.

Owens, J. and Hunter, A. (2000). Application of the Self-organizing Map to Trajectory Classification, *In Proceedings of IEEE International Workshop on Visual Surveillance*, 77-83.

Papadakis, V. M., Papadakis, I. E., Lamprianidou, F., Glaroulos, A., and Kentouri, M. (2012). A Computer Vision System and Methodology for the Analysis of Fish Behavior, *Aquacultural Engineering*, 46, 53-59.

Piciarelli, C., Micheloni, C., and Foresti, G. L. (2008). Trajectory-based Anomalous Event Detection, *IEEE Transactions on Circuits and Systems for Video Technology*, 18(11), 1544-1554.

Pinkiewicz, T. H., Purser, G. J., and Williams, R. N. (2011). A computer vision system to analyze the swimming behavior of farmed fish in commercial aquaculture facilities: A case study using cage held atlantic salmon, *Aquacultural Engineering*, 45(1), 20–27.

Porikli, F., and Haga, T. (2004). Event Detection by Eigenvector Decomposition Using Object and Frame Features, *In Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR)*, 114-122.

Pudil, P., Novovicova, J., and Kittler, J. (1994). Floating Search Methods in Feature Selection, *Pattern Recognition Letters,* 15(11), 1119-1125.

Schalie W.H., Shedd T.R., Knechtges P.L., and Widder M.W. (2001). Using higher organisms in biological early warning systems for real-time toxicity detection. *Biosensors and Bioelectronics*, 16(7), 457–465.

Serra-Toro C., Montoliu, R., Traver, V. J., and Hurtado-Melgar, I. M. (2010). Assessing Water Quality by Video Monitoring Fish Swimming Behaviour, *In Proceedings of IEEE International Conference on Pattern Recognition (ICPR)*, 428-431.

Silla C.N., Freitas A. A. (2010). A survey of hierarchical classification across different application domains, *Data Mining and Knowledge Discovery*, 22(1-2), 31–72.

Silla C. N. and Freitas A. A. (2011). Selecting different protein representations and classification algorithms in hierarchical protein function prediction, *Intelligent Data Analysis Journal*, 15(6) 979-999 2011.

Sillito, R. R., and Fisher, R. B. (2008) Semi-supervised Learning for Anomalous Trajectory Detection, *In Proceedings of British Machine Vision Conference (BMVC),* 227-238.

Spampinato, C., Giordano, D., Salvo, R. D., Chen-Burger, Y., Fisher, R. B., Nadarajan, G. (2010). Automatic Fish Classification for Underwater Species Behavior Understanding, *In Proceedings of First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams*, 45-50.

Spampinato, C., Palazzo, S., Giordano, D., Kavasidis, I., Lin, F-P, Lin, Y-T. (2012). Covariance Based Fish Tracking In Real-Life Underwater Environment, *In Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP)*, 2012.

Thida M., Eng H., and Chew B. F. (2009). Automatic analysis of fish behaviors and abnormality detection. *In Proceedings of International Association for Pattern Recognition Conference on Machine Vision Applications*, 8(18), 278-282.

Xiang T. and Gong S. (2005). Video behaviour abnormality detection using reliability measure. In Proceedings of British Machine Vision Conference, 2005.

Xiang, T., and Gong, S. (2006). Incremental Visual Behaviour Modelling, *IEEE Visual Surveillance Workshop*, 65-72.

Xu, J., Liu, Y., Cui, S., and Miao, X. (2006). Behavioral responses of tilapia (oreochromis niloticus) to acute fluctuations in dissolved oxygen levels as monitored by computer vision, *Aquacultural Engineering*, 35(3), 207–217.

Xu, X., Tang, J., Liu, X., and Zhang, X. (2010). Human Behavior Understanding for Video Surveillance: Recent Advance, *In Proceedings of International Conference on Systems Man and Cybernetics*, 3867-3873.

Zhang, D., Gatica-Prez, D., Bengio, S., and McCowan, I. (2005). Semi-supervised Adapted HMMs for Unusual Event Detection, *In Proceedings of IEEE Computer Vision Pattern Recognition (CVPR)*, 1, 611-618.

Zhong, H., Shi, J., and Visontai, M. (2004). Detecting Unusual Activity in Video, *In Proceedings of Computer Vision Pattern Recognition (CVPR)*, 2, 819–826.

Table 1

Possible class decisions for a new trajectory

Decision	Condition	Next Action
Unusual Trajectory	• The closest cluster is a perfectly classified pure unusual cluster or • The closest cluster is a perfectly classified mixed cluster and the new trajectory is further than the outlier detection threshold of that cluster or • The closest cluster is a perfectly classified pure normal cluster and the new trajectory is further than the outlier detection threshold of that	Classification stops (there is no need to look at any other level of the hierarchy).
Candidate Normal Trajectory	• The closest cluster is a perfectly classified pure normal cluster and the distance between the new trajectory and the corresponding cluster's center is smaller than the outlier detection threshold of that cluster or • The closest cluster is a perfectly classified mixed cluster and the distance between the new trajectory and cluster center is smaller than the threshold.	The new trajectory goes to the next hierarchy level.
No Decision	• The closest cluster is a misclassified cluster.	The new trajectory proceeds to the next level.

Table 2

Class decisions for a new trajectory when the closest cluster at each level is a misclassified cluster.

Decision	Condition	Next Action
Unusual Trajectory	• The closest cluster at the current level contains all normal trajectories by looking at the ground-truth class labels and the new trajectory is further than the rest of the samples in that cluster or • The closest cluster contains all unusual training trajectories by the ground-truth or • The closest cluster contains both normal and unusual training trajectories. Then, the nearest neighbor rule which makes the class of the new trajectory the same as the closest training sample's class is applied. If the class is an unusual class then this decision and corresponding action is given.	Classification stops (there is no need to look at any other level of the hierarchy).

Candidate Normal Trajectory	• The closest cluster at the current level contains all normal trajectories by looking at the ground-truth class labels and the new trajectory is not further than the rest of the samples in that cluster or • The closest cluster contains both normal and unusual training trajectories. Then, the nearest neighbor rule which makes the class of the new trajectory the same as the closest training sample's class is applied. If the class is an unusual class then this decision and corresponding action is given.	The new trajectory goes to the next hierarchy level.
Normal Trajectory	• If the new trajectory reaches the last level and could not be classified yet.	Classifications stops.

Table 3

Best average GeoMean result of each method with the corresponding TPrate and *TNrate using the fish trajectory dataset. The best results of each metric are emphasized in bold-face.*

Methods	TPrate	TNrate	GeoMean
KNN	0.26±0.08	**0.99±0.01**	0.50±0.09
KNN-wFS	0.37±0.28	**0.99±0.01**	0.60±0.27
SVM	0.21±0.07	**0.99±0.01**	0.45±0.07
SVM-wFS	0.81±0.16	0.93±0.03	0.86±0.09
RF-BT	0.87±0.01	0.93±0.06	0.90±0.03
RF-BT-wFS	0.88±0.01	0.91±0.10	0.89±0.05
UMOT	0.57±0.2	0.85±0.11	0.70±0.04
LOF	0.62±0.17	0.97±0.01	0.77±0.08
Filtering	0.80±0.20	0.77±0.04	0.78±0.09
FlatClass	0.81±0.17	0.76±0.02	0.78±0.09
Proposed	**0.94±0.10**	0.88±0.02	**0.91±0.05**

Table 4

Best average GeoMean result of each method with the corresponding TPrate and *TNrate using the Forum Pedestrian Database (Majecka, 2009). The best results of each metric are emphasized in bold-face.*

Methods	TPrate	TNrate	GeoMean
SVM-wFea	0.83±0.03	0.79±0.04	0.81±0.01
RF-BT	0.80±0.02	0.86±0.03	0.83±0.02
RF-BT-wFea	0.79±0.04	0.81±0.05	0.80±0.04
LOF	0.53±0.07	**0.95±0.02**	0.71±0.04
Proposed	**0.87±0.06**	0.86±0.05	**0.86±0.02**

Table 5

Best best average GeoMean results of given methods in Table 4 with corresponding TPrate and *TNrate using the fish trajectory dataset. The best results are emphasized in bold-face.*

Methods	TPrate	TNrate	GeoMean
Proposed	**0.94±0.10**	0.88±0.02	**0.91±0.05**
SingleLevProposed	0.58±0.16	0.90±0.03	0.72±0.10
Alter1	0.37±0.16	0.97±0.01	0.59±0.13
Alter2	0.92±0.02	0.80±0.17	0.85±0.09
Alter3	0.88±0.10	0.91±0.02	0.89±0.05
Alter4	0.48±0.21	0.96±0.02	0.68±0.17
SVM-wPropFea	0.89±0.11	0.86±0.05	0.87±0.06
Hie-SVM	0.92±0.10	0.82±0.09	0.86±0.02
Hie-SVM-Alter1	0.36±0.34	**0.98±0.03**	0.59±0.34

Figure Captions

Figure 1 Hierarchy Construction

Figure 2 New Trajectory Classification using the Constructed Hierarchy

Figure 3 Example of (a) normal fish trajectory, (b) unusual fish trajectory.