The Identification Process: Human and Machine (AVEDac) Efficiency and Effectiveness

Samantha Peterson, Florida Institute of Technology, Melbourne, Florida

Mentor: Danelle Cline

Summer 2012

Keywords: AVEDac, Event Detection, Self Consistency, Plankton, Holothurians

ABSTRACT

The need to process large amounts of video and image data in the marine sciences fields has greatly increased with the development of Remotely Operated Vehicles (ROVs) and other submersible recording devices. This need has pushed researchers and scientists to develop programs that aid in detection and classification of events within these videos and images known as the identification process, like the Automated Visual Event Detection and Classification (AVEDac) program at MBARI. This study aims to identify the strengths and weaknesses within the human identification process and the machine process by detecting and classifying plankton images from LISST-HOLO data and holothurian videos taken with one of MBARIs ROVs. We found that both humans and machine have difficulties with identifying organisms to genus and species level. Also, humans had a higher ability to detect organisms than AVEDac produced. This was probably caused by difficulty of separating holothurian color and texture form that of the benthic floor due to their similarities. This test was a preliminary test and future studies must be done in order to collect valuable data to improve the identification process for both humans and machine.

INTRODUCTION

Over the years, the use of Remotely Operated Vehicles (ROVs) has given scientists and researchers within the marine science fields an influx of raw data documenting life underwater. Videos can now be recorded documenting interactions, populations, and new species in a world that was impossible to see and observe beforehand. Through the use of ROVs, scientists can document these things, as well as collect samples that will aid in research and education. This influx of data is a valuable source but processing this large quantity poses a problem for researchers. Hours and hours of video are captured on ROVs, like the Ventana and Doc Rickets and the Monterey Bay Aquarium Research Institute (MBARI), waiting in their extensive libraries to be processed and analyzed.

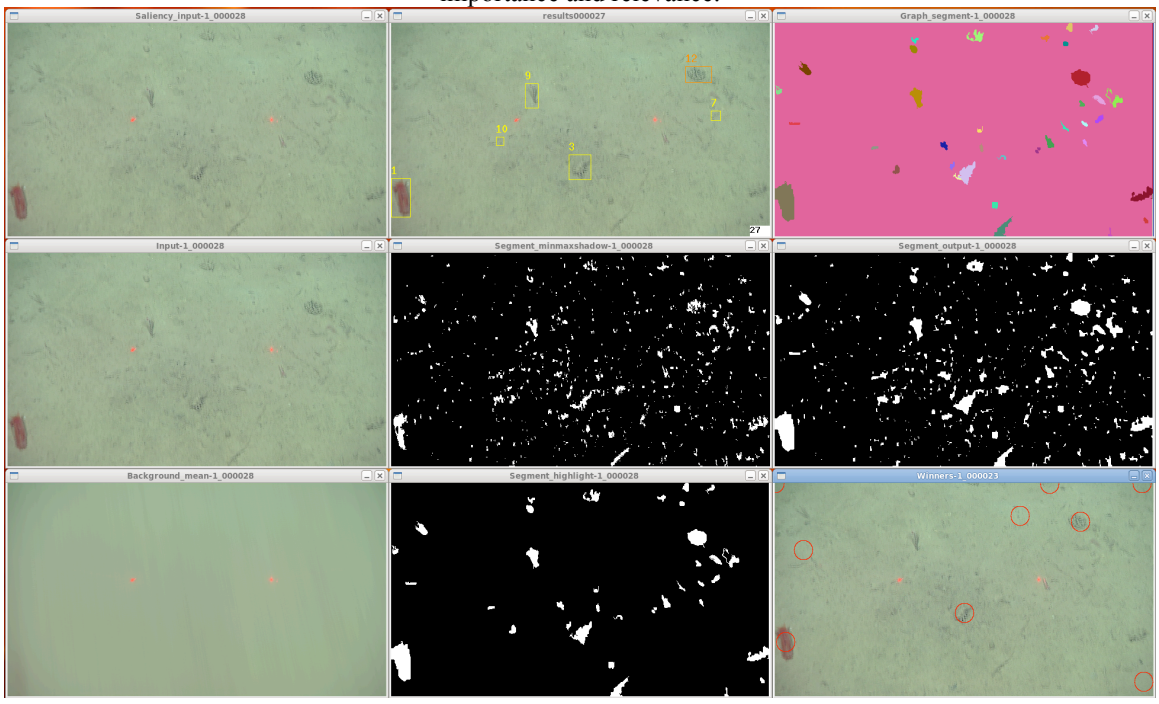
Researchers at MBARI use the Video Annotation and Reference System, also known as VARS, to aid in the identification and analysis of these vast quantities of videos. VARS works as a reference tool, storing more than 3,500 terms describing organisms, substrate, and technical terms used to annotate the hundreds of hours of video captured from ROVs and other underwater documentation (Schilining, 2006). As lab technicians watch these videos, they record events and interesting images. This process is one of the best that there is but it still faces challenges. For example, though the images are annotated, the annotations do not tell you the position of the animal within the frame, this causes problems when there are multiple organisms within a screen. Also, this process is still extremely time consuming: Lab technicians must go through hours and hours of empty recordings to find interesting events and the amount of data to process is never ending.

Even with this aid in the identification process, the question of how well humans can identify organisms and their consistency in doing so is one that is constantly asked. Researchers have reported that humans have a “superb visual perception” when identifying objects but distractions like fatigue, boredom, eye strain, and other physical ailments , as well as short term memory decay over time, affect their ability to correctly identify organisms (Culverhouse, 2007). In study to compare the identification of dinoflagellates by humans and machine, it was found that subjects have trouble

identifying separate species within the genus *Dinophysis* due to similarities in structure (Culverhouse, 2003), proving the difficulties nature forces upon us.

These reasons were some of the drivers in the development of the Automated Visual Event Detection and Classification (AVEDac) system by engineers at MBARI. AVEDac analyses images and videos recorded by ROVs, as well as Autonomous Underwater Vehicles (AUVs) and other oceanic devices, to identify organisms and events of interest. The program works like the human eye, looking at frames of the videos and still images to find what it determines as "interesting" by using texture, color, and other factors. AVEDac can be programmed to search for and identify larger organisms, such as fish, as well as organisms as small as plankton. Figure 1 shows the different aspects that AVEDac analyzes in order to determine interesting events, such as screen segmentation, shadows, movement or event, saliency, and length of time "event" lasts as something interesting, the end result being an event captured to be identified. The figure also shows you that the processing does pick up on junk that it determines as "interesting". This can be changed by the size AVEDac is programmed to range between for event detection as well as the length of time that the program stared at the screen; these "junk events" can also be deleted or kept for improved calibration.

Figure 1. Depiction of the layers AVEDac uses to identify interesting events and determine their importance and relevance.



This study aims to test the relationship between human and machine identification, using the AVEDac program. There are two aspects to the study that differ greatly. The first aims to establish the consistency of classification in human test subjects that are not very experienced in the identification field. To do this, images taken from the LISST-HOLO analysis would be identified. The second aspect aims to compare detection and classification between human subjects and AVEDac by identifying Holothurians, or sea cucumbers, on the sea floor. The purpose of these tests is to establish the weaknesses in both identification methods and find ways of improvement in the future.

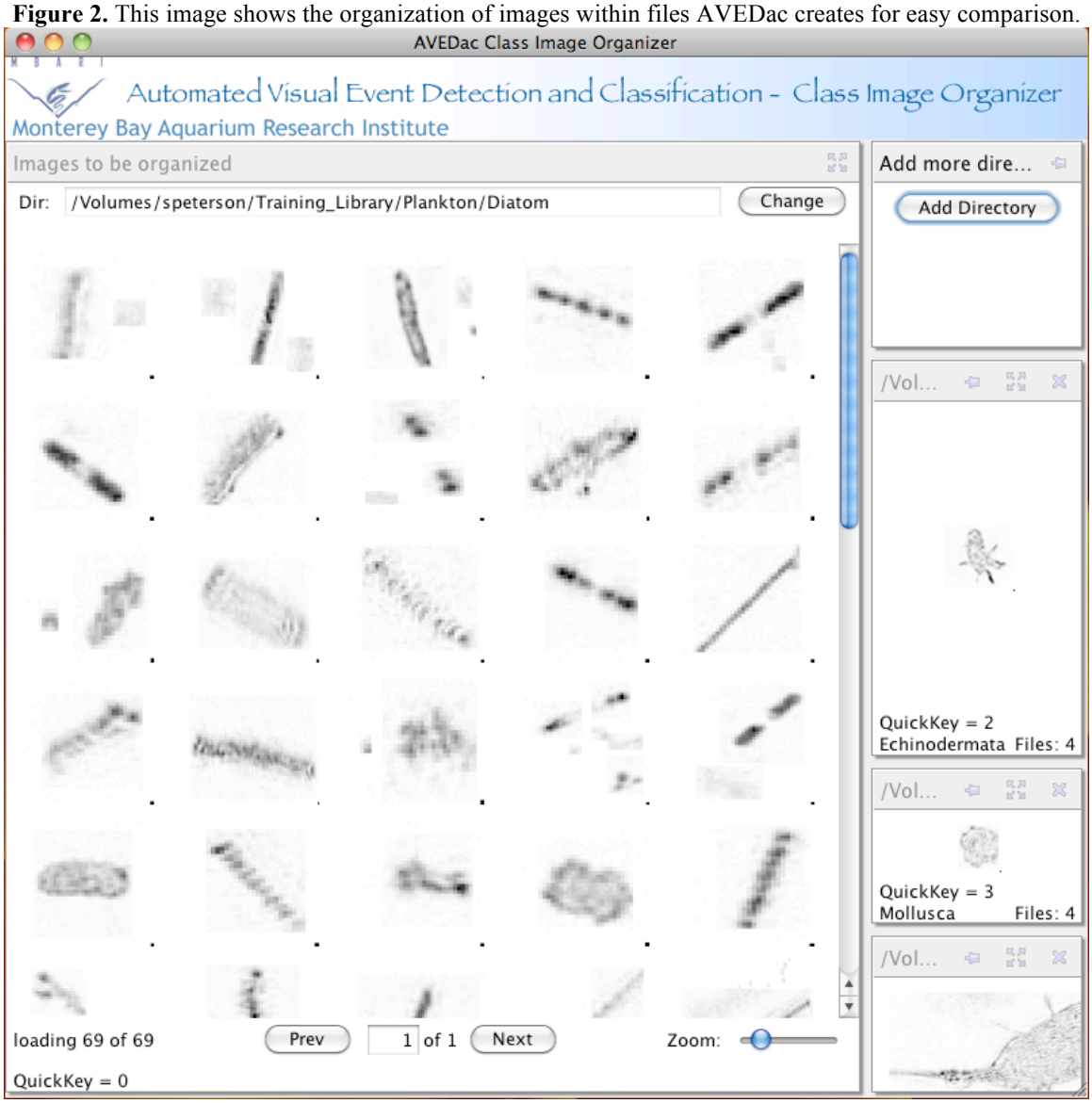
MATERIALS AND METHODS

A) HUMAN CONSISTENCY OF PLANKTON CLASSIFICATION

Processing Raw Data In order to test for consistency, LISST-HOLO images of plankton were chosen as the testing medium. These samples were taken in Monterey Bay in early 2012 by research specialist, John Ryan, and his team at MBARI. The data had not been processed before being used for this experiment so the images, which totaled about 14,000, were split into files containing 100 images and made into a video. This video was then processed through AVEDac to identify "interesting" events. Due to the size of the images and the fact that these "videos" were compressed still images, AVEDac was programmed to pick up extremely small events and stared at each frame for a longer period of time than that of a video being processed.

Creating an Image Directory In order to have a pool of images to choose from, a image directory was created separating images into categories of organisms. This was done by manually viewing each file, sorting out the viable images and removing the rest, and then labeling each image with an organism. The list of organisms included: Diatom, Dinoflagellate, Rotifera, Echinodermata, Copepoda, Crustacean, Mollusca, Polychaeta, Bryozoa, and Unknown, as well as a Junk category for images that were not suited for any of the previous categories. Figure 2 shows an example of an Image Organizer that AVEDac makes throughout the process. Of the 14,000 images available for processing, 2,700 were manually processed and labeled to build these libraries. The largest file of organized images, in respect to organisms, was the Diatoms with 69 images.

Figure 2. This image shows the organization of images within files AVEDac creates for easy comparison.



Building Test Images Once a viable file was created for each organism, three files were created with ten manually selected images within each, representing each category, excluding "Unknown." Each file had ten images that were shuffled and unlabeled so that there was no bias for the testers in the classification process. Table 1 shows the master list of organisms within each file. Organisms were repeated to determine the consistency of testers.

	File 1	File 2	File 3
1	Diatom (Chaetoceros)	Dinoflagellate	Copepod
2	Echinodermata	Crustacean	Diatom (Detontula)
3	Diatom (Stephanopyxis)	Copepod	Bryozoa
4	Rotifera	Dinoflagellate	Junk
5	Junk	Mollusca	Dinoflagellate
6	Diatom (Detonula)	Diatom (Stephanopyxis)	Mollusca
7	Polychaeta	Echinodermata	Diatom (Chaetoceros)
8	Crustacean	Junk	Crustacean
9	Dinoflagellate	Diatom (Detontula)	Diatom (Stephanopyxis)
10	Copepod	Rotifera	Echinodermata

Table 1. Master list of organisms within each testing file.

Testing Four individuals were chosen to test for consistency. Each individual possessed a different level of experience, ranging from no experience to mild experience it the identification and classification of various organisms. Each test subject was given a resources library to give a basis of possible organisms within the files, which was built using the Plankton Identification website provided by CeNCOOS & HABMAP (CeNCOOS) and "A Guide to the Marine Plankton of Southern California 3rd Edition" provided by UCLA OceanGLOBE (Perry 2003). This resource can be found in Appendix 1. Each file was opened in AVEDac one at a time to be identified. Time taken to identify images was documented for later comparison. After each file was finished, testers had a 20 minute break in order to ensure enough time separating each classification and reset their memory. This break insured that testers were not identifying species off of memories from the former file, but learning as they went and associating the resource with the organism within the image.

Analysis Identified files were then exported into an excel spreadsheet to calculate correctness and consistency. This was done simply by counting all labeled images that were correct when compared to the master list and then graphed to show differences.

B) HUMAN AND MACHINE: DETECTION AND CLASSIFICATION OF HOLOTHURIANS

Processing Raw Data A ten minute clip of a dive taken in last 2011, abundant in holothurians was used for this comparison test. The video contained the following holothurians species: Amperima-Peniagone Complex, Peniagone sp., Elpidia sp., and *Scotoplanes globosa*, as well as other organisms that were not of interest. The ten minute clip was split into two videos, with five minutes each: Avedbenthic2012-1.mov and Avedbenthic2012-2.mov. These videos were then processed in AVEDac to identify interesting events, mainly holothurians, in a similar way to the LISST-HOLO processing. Due to the fact that this resource was a video instead of still images, the program was programmed to stare at each frame for a shorter period of time in comparison to the time spent on LISST-HOLO images. This data was also used in comparison with human identification.

Creating Training Libraries Three training libraries were created for the purpose of this experiment. Each library was built in the same way as the training libraries for the LISST-HOLO images. Labels were different in each library and were of different sizes. Libraries and their labels are as follows:

Holothurian #1	
Amperima-Peniagone Complex Peniagone sp. B Peniagone sp. C Peniagone sp. D Elpidia *Scotoplanes globosa* *(only 1st video)*	**Holothurian #4**
	Amperima-Peniagone Complex Peniagone Elpidia *Scotoplanes globosa* *(both 1st and 2nd video)*
Holothurian #2	**Holothurian #5**
Amperima-Peniagone Complex Elpidia *Scotoplanes globosa* *(only 1st video)*	Amperima-Peniagone Complex Peniagone sp. A Peniagone sp. B Peniagone sp. C Peniagone sp. D Elpidia *Scotoplanes globosa* *(both 1st and 2nd video)*
Holothurian #3	
Amperima-Peniagone Complex Elpidia *Scotoplanes globosa* *(both 1st & 2nd video)*	

Training libraries #1-2 had images that were only from events within file Avedbenthic2012-1.events.xml, while training libraries #3-5 had images from both the first and second video (files Avedbenthic2012-1.events.xml and Avedbenthic2012-2.events.xml) for separate processing. Individual species were identified with the use of a master list provided by an expert in holothurians identification, Linda Kuhnz.

Human Testing Three subjects were chosen to test detection and classification of holothurians, each had different levels of experience (none, some, and a lot) in identifying invertebrates. These were three of the same individuals that participated in the plankton classification. They were given a resource manual with images and descriptions of holothurians present in the videos being processed, built from the online MBARI Deep Sea Guide (MBARI, 2012), which can be found in Appendix 2. The individuals were tested on their ability to identify holothurians presence, how long it took them to identify individuals, and their accuracy of classifying holothurians species within the second of the two five minute clips.

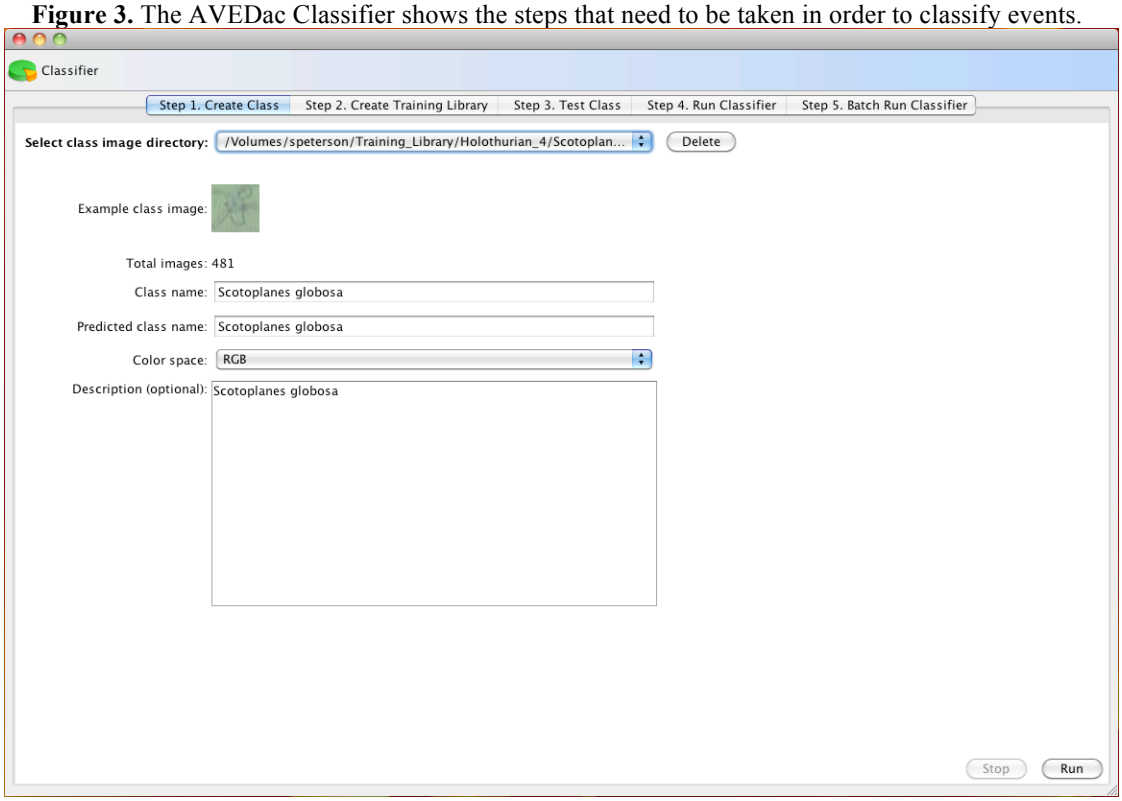
Video Detection Individuals were instructed to observe holothurians within file Avedbenthic2012-1.mov and document the frame number and position of the organisms throughout the video. This documentation was then recorded in Excel and compared with the master list of frame number occurrences developed before detection. The detections were then compared by percentage identified and percentage missed.

Classification of Species After detection, individuals were instructed to classify each "interesting" event picked up by AVEDac by species of holothurians within file Avedbenthic2012-1.events.xml. To minimize confusion, when preparing the file, all events that were not of holothurians were deleted. This was timed as well, to compare the ease of classification later. They used Holothurian Training Library #1 & #2 classification labels. Each individual's labels were then exported to Excel and compared to the master list and calculated for percentage correct and percentage incorrect.

AVEDac Program Identification The AVEDac classification process takes multiple steps to prepare after the training libraries have been built. The AVEDac

Classifier was used to test the accuracy of AVEDac in classification using each of the libraries on the events of the second video (Avedbenthic2012-2.events.xml). The steps to create a library and produce AVEDac Classification are shown in Figure 3: Create Class, Create Training Library, Test Class, and Run Classifier were used. Step 5: Batch Run Classifier is used when multiple files are being tested, which was not applicable in this experiment.

Figure 3. The AVEDac Classifier shows the steps that need to be taken in order to classify events.



Creating Classes Each Image Directory is sorted into its own file and compressed into a "class" by collecting all of the images within the file. These are then remembered by AVEDac in order to make a Training Library.

Creating Training Libraries Training libraries were then created by selecting Classes that were desirable for the test. Multiple Training Libraries were created in order to determine accuracy of classification. These training libraries were tested against classes in "Step 3: Test class" by programming the library to identify the images within the class selected. This helped determine if the library was accurate and useful for texting.

Running the Classifier Once the Training Libraries were built, they were used to run the classifier against the events file for the second video. This compared the manually classified labels, from the master list provided by the expert, of the events to the library. Once compared, AVEDac "predicted" which class the events would fall into, in the process AVEDac would develop an "Unknown" classification if it was unable to place the event into one of the classes. Along with predicting a class, AVEDac also produced a Confusion Matrix outlining what was labeled correctly and what was labeled incorrectly when compared to the manually classified labels, recognized by AVEDac as "actual" organism classifications. These were run at 70%, 80%, and 90% probability to produce a ROC (Receiver Operating Characteristic) Curve for each training library. This changed the threshold that the images had to fall within to be labeled a certain class.

Analyzing The predicted labels and confusion matrixes were exported to Excel to be organized for MATLAB. Predicted Labels were given numbers that corresponded to each class and run through MATLAB to produce True and False Positives (TP & FP), and True and False Negative (TN & FN), as well as the Sensitivity. With this output, ROC Curves were produced to show the trend and accuracy of AVEDac's classification. These were then compared to the percentages found from the Human identifications.

C) CONCLUDING SURVEY

Each participant was given a short survey, found in Appendix 3, asking what key aspects they looked at for detection and classification, what the challenges were, what was useful, what would be useful in the future and other questions. This information was used to compare what they used to determine identification in the end of the experiment.

RESULTS

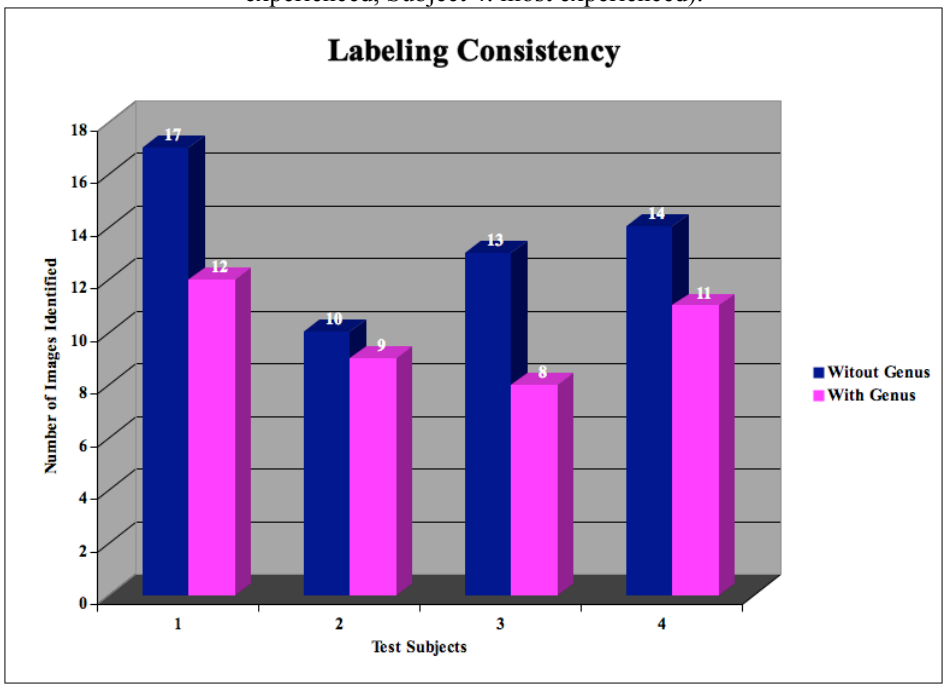
A) HUMAN CONSISTENCY OF PLANKTON CLASSIFICATION

The average time needed to identify a file of ten images when using the reference guide was about 10.5 minutes. The longest length of time needed to identify a single file

was 23 minutes by Subject 1 in the first file of images, while the shortest was 5 minutes by Subject 2 in the third file of images.

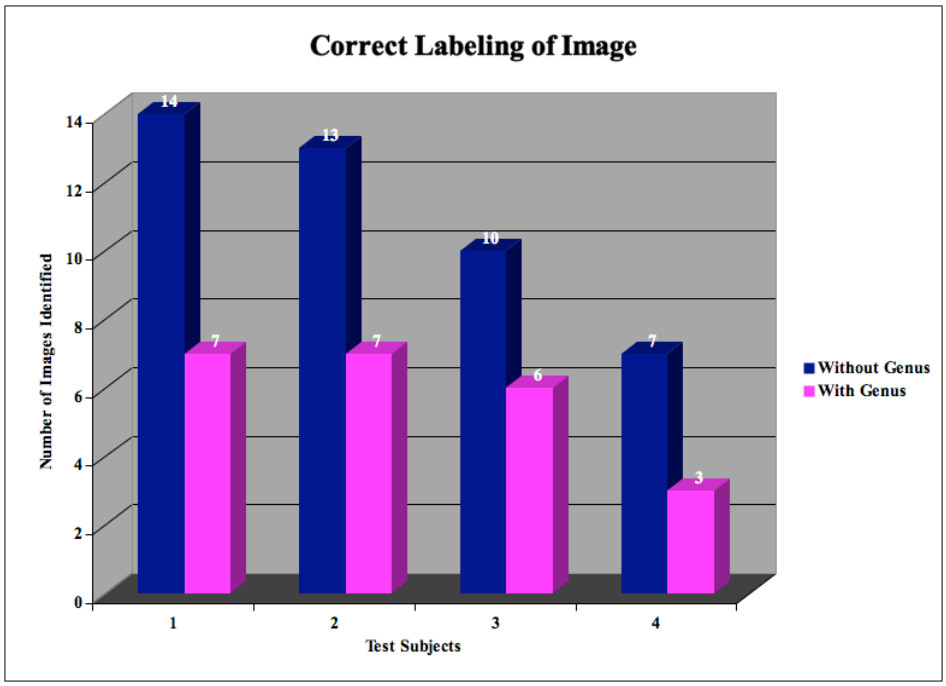
The data from the four test subjects on plankton classification showed less than a 50% consistency in classification of organisms, as shown in Figure 4. Subject 1, who had the least experience in identification within the group, had the greatest consistency value when not taking genus classification into effect, with 17 out of 30 identified with constancy. Subject 1 also had the highest consistency when taking into effect the genus of species with 12 out of 30, though it dropped.

Figure 4. Graph shows little trend association between skill level and consistency (Subject 1: least experienced, Subject 4: most experienced).



The data was also compared for correct labeling. The highest number of images labeled correctly when not taking genus into effect, according to the master list, was 14 out of 30 by Test Subject #1. The lowest number of images labeled correctly for both without genus and with genus consideration, was from Test Subject #4, who was the most experienced test subject in previous identification of various organisms, as shown in Figure 5.

Figure 5. This graph shows a decrease in accuracy in relation to experience of identification increase.



Both accuracy and consistency decreased when genus of Diatoms and Dinoflagellates were accounted for. When it came to consistency it was not as effective, however, in accuracy, it cut most of the values in half.

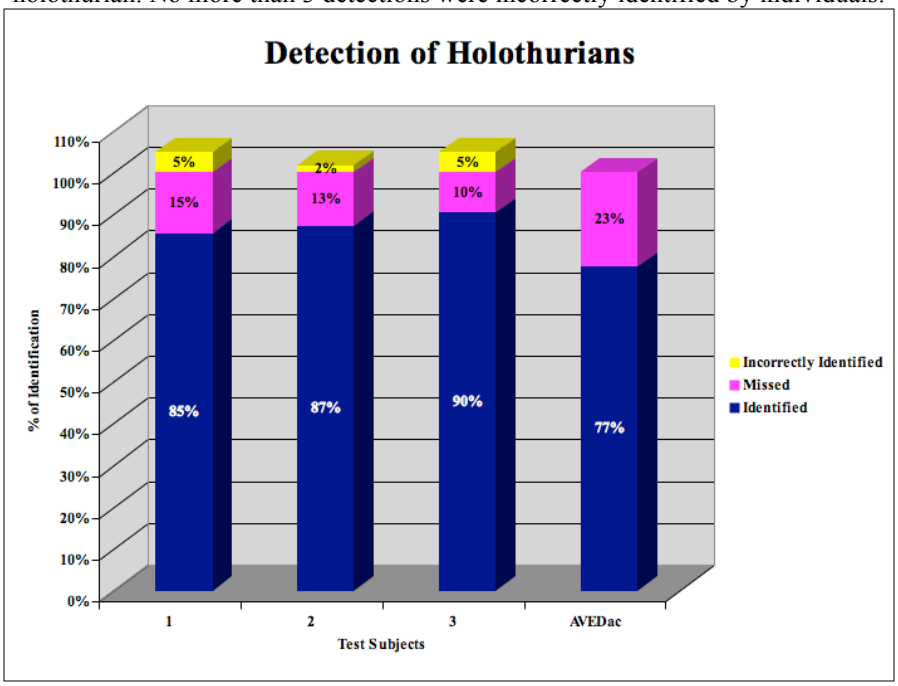
B) HUMAN AND MACHINE: DETECTION AND CLASSIFICATION OF HOLOTHURIANS

Human and AVEDac Detection The average length of time it took for testers to identify holothurians within the five-minute video was 23 minutes, with the highest time being 27 and the lowest being 20. There is no data on how long it took AVEDac to isolate only holothurians when being processed, though the entire processing took over 12 hours to produce the entire file of events.

When comparing human and AVEDac detection to the total known number of holothurians within the video, humans scored a higher percentage: ranging from 85-90%, while AVEDac produced 77% detection (Figure 6). There were a total of 62 holothurians present in the video. The highest number recorded by testers was 56 individual holothurians by Subject 3.

Figure 6. Percentage breakdown of identified and missed holothurians shows humans had the better ability to identify most of the holothurians present in the video. The yellow in the graph shows the percentage of

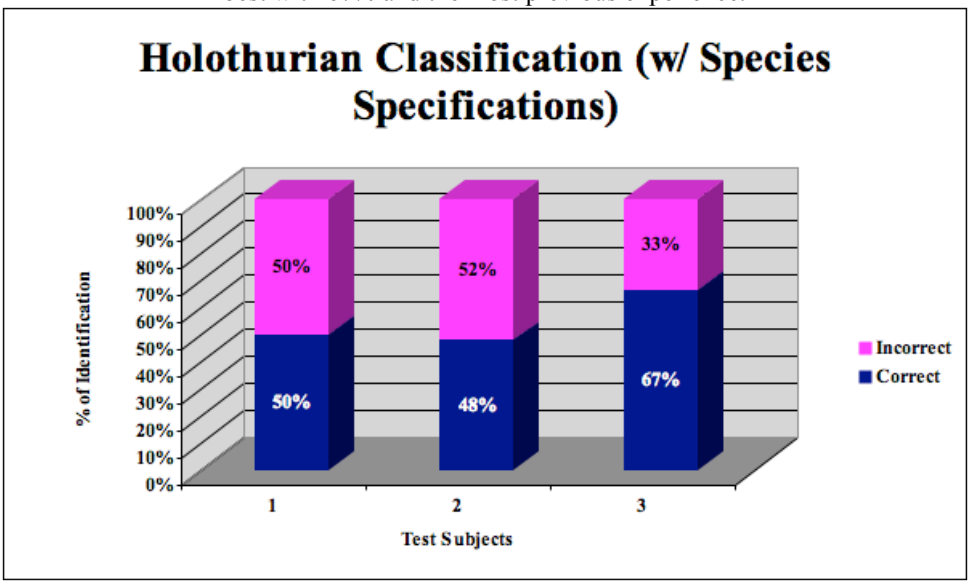
incorrectly identified holothurians, organisms or interesting things that the tester believed to be a holothurian. No more than 3 detections were incorrectly identified by individuals.



Human Holothurian Species Classification On average, it took about 24.6 minutes for testers to identify the species of the 48 holothurians put into "events" with AVEDac. The longest time was 29 minutes, while the shortest was 22 minutes.

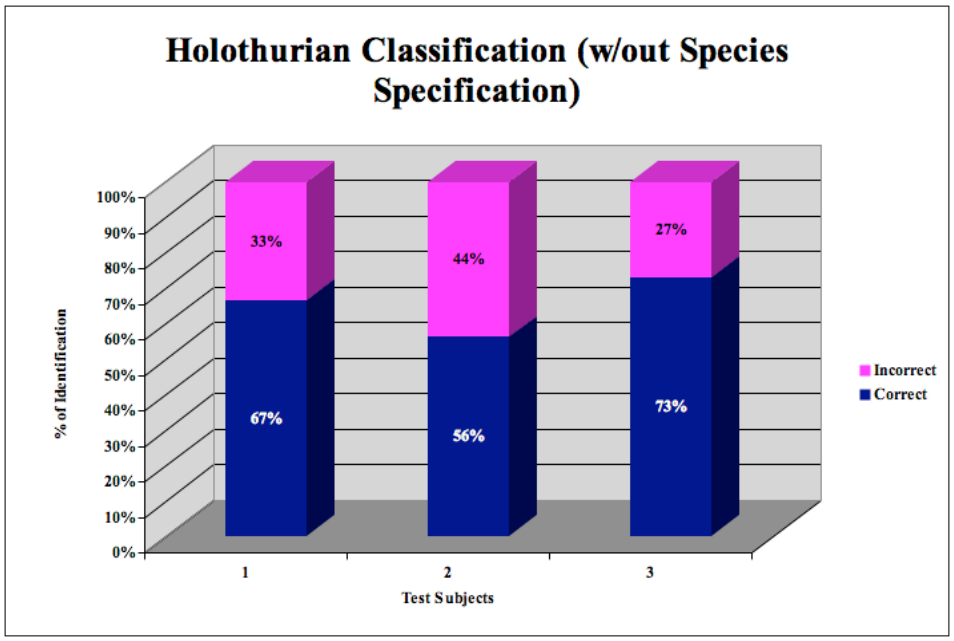
Classification with Species Specifications Testers used the following labels to identify species: Amperima-Peniagone Complex, Peniagone sp. A, Peniagone sp. B, Peniagone sp. C, Peniagone sp. D, Elpidia, and *Scotoplanes globosa*. With these class labels, the highest percentage correctly labeled by testers was 67% and the lowest being 48% (Figure 7). The most mislabeled events were those within the Amperima-Peniagone Complex and the Peniagone species. Few Elpidia and *Scotoplanes globosa* events were confused as well.

Figure 7. Testers showed a range of 48-67% accuracy of species classification. Subject 3 performed the best with 67% and the most previous experience.



Classification without Species Specifications Due to similarities, the Peniagone species choices were combined to test for accuracy when excluding species breakdown. The results showed that accuracy increased by 6-17% for each tester. Instead of accuracy being at 67% as a high, it increased to 73% (Figure 8).

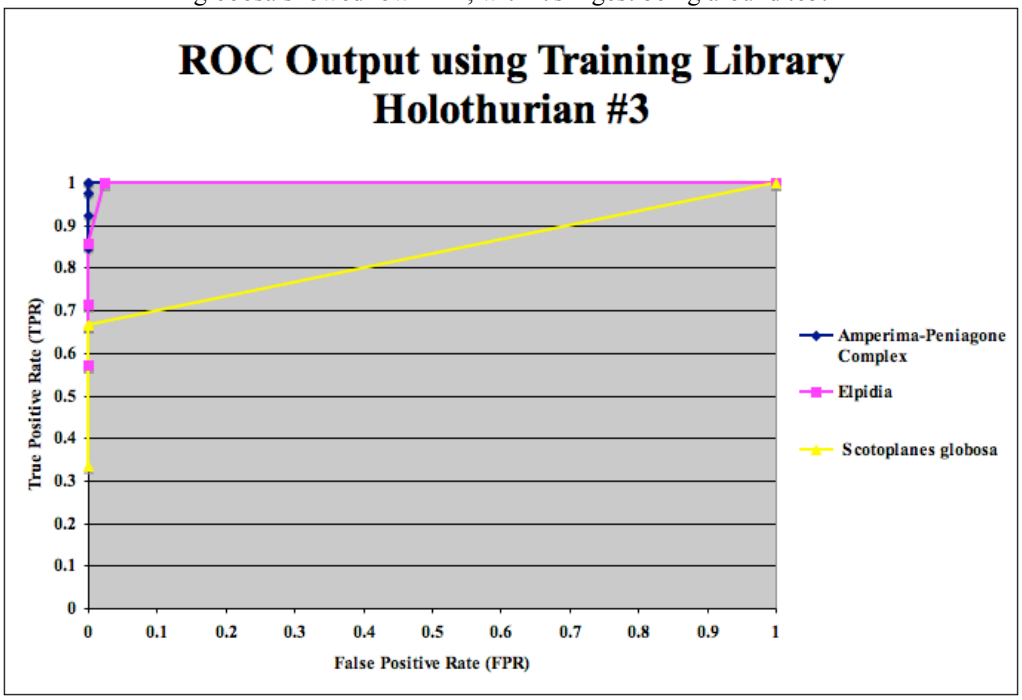
Figure 8. Correct labeling increased between 6-17% when Peniagone species were combined, due to their similarities.



AVEDac Holothurian Species Classification The accuracy of AVEDs classification effort was tested in multiple parameters in order to identify problems and to see differences. Training Libraries (TR) Holothurian 3, 4, & 5 were used to create confusion matrixes and ROC (Receiver Operating Characteristic) curves to determine accuracy. They were analyzed within the AVEDac Classifier and tested at multiple probability thresholds, using the RGB color choice. The data from the confusion matrix produces and the predicted labels were then used in MATLAB to produce the ROC curves. The reported results will start with the simplest breakdown of classes and then move up to the more complicated, broken down classes that are sectioned into specific species classifications.

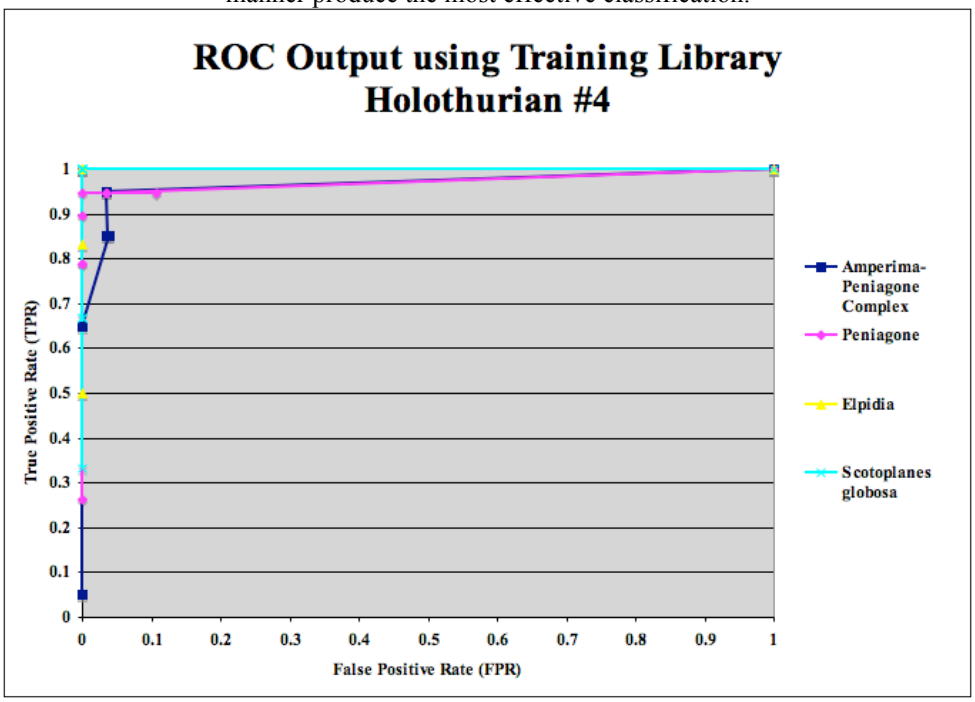
Training Library Holothurian #3 This library contained the following classes: Amperima-Peniagone Complex, Elpidia, and *Scotoplanes globosa.* The probability thresholds were tested at 40%, 50%, 60%, 70%, 80%, 90%, and 99% and plotted on ROC curves. The ROC curves for TR Holothurian #3 indicate that the combined Amperima-Peniagone Complex had a high True Positive Rate (TPR) of identification, this indicated that most of the events within this class were identified correctly as probability thresholds increased.

Figure 9. When the Peniagone species were combined within the Amperima-Peniagone Complex, dude to their similar morphologies, they show a high True Positive Rate (TPR) of identification. Scotoplanes globosa showed low TPR, with it's higest being around .65.



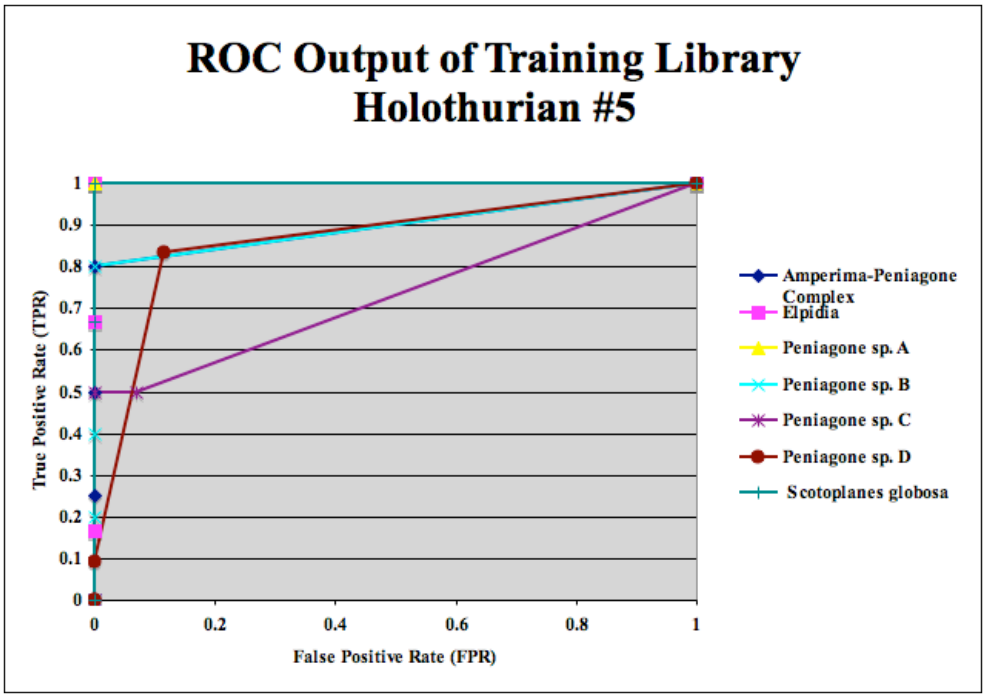
Training Library Holothurian #4 TL Holothurian #4 was tested at 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 99% probability thresholds of classes Amperima-Peniagone Complex, Peniagone, Elpidia, and *Scotoplanes globosa*. The ROC curves improved for all classes with the additional separation of the Peniagone events. All showed high TRP values, with extremely low False Positive Rate (FPR) values, indicating accurate classifications (Figure 10).

Figure 10. The ROC curves for TR Holothurian #4 indicate that the separation of the holothurians in this manner produce the most effective classification.



Training Library Holothurian #5 Training Library #5 had the most broken down classification classes which showed to have a large effect on the ROC curves produces, shown in Figure 11. Due to the breakdown of similar species, the curves were not accurate in classification. Peniagone sp C. had the worst TPR values, with the highest value at 0.5, corresponding with a FPR of 0.069. On the other end, the Amperima-Peniagone Complex, Elpidia, and Peniagone sp. A had TPR values of 1 and FPR values of 0 across all probability thresholds. In the case of Peniagone sp. A, there was only one event to be classified as within this class, showing that the classifier classified this event correctly at each probability threshold.

Figure 11. The separation of species shows to effect the overall classification of all classes, indicating interference between them due to similarities and difficulties.



Testers Survey Results Overall, testers reported that plankton classification was considerably more difficult than holothurian detection and classification, with an average difficulty rating of 7.75 on a scale of 1-10 (10 being the most difficult) compared to the average of 4.83 for the holothurian process. They also reported that the resources were helpful during classification, but the plankton resource could be improved with images closer in similarities to the LISST-HOLO images.

Shape was the highest used and important factor to all participants in the classification of both plankton and holothurians, with color being the next important factor for holothurian classification. All participants reported that the quality of the images and video would have improved their classifications as well, though the video quality was considerably better than that of the plankton still images.

DISCUSSION

A) HUMAN CONSISTENCY OF PLANKTON CLASSIFICATION

The results suggest that there is little consistency in human classification of plankton organisms. Though there were exceptions, such as Echinodermata having the

highest accuracy rate and consistency rate of classification out of all organisms present, the overall conclusion from this part of the study was that consistency was not strong.

This inconsistency could be due to the fact that the testers were not experienced identifiers. Plankton are very difficult to identify and are often only correctly identified when the identifier is highly trained in their field, usually only with specific genus or species (Culverhouse, 2007). This is due to plankton's subtle differences, which make them challenging to identify. The fact that testers had higher accuracy and consistency when genus was not taken into account also enforces the idea that the subtle differences of species make them difficult to identify and can be misinterpreted easily.

In the future, testers that are used should be more experienced in the ability to identify plankton. The analysis of LISST-HOLO images will be useful in the future, but only if experienced identifiers are used. Their classifications will aid in accurately classifying genus in future research needs and tests. Also, the test size of images to identify should be increased to get a wider range of consistency instead of having 3 points of similarity. Increased number of testers would also improve data to possibly show a trend in experience levels that is more consistent, as well as provide more data to establish a average accuracy rate.

B) HUMAN AND MACHINE: DETECTION AND CLASSIFICATION OF HOLOTHURIANS

Detection In this experiment, humans showed to have stronger abilities to detect organisms, in this case: holothurians, when compared to the abilities of AVEDac. Testers scored higher across the board than the detection rate of AVEDac.

A possible improvement to AVEDac's detection rate would be to allow the program to process the video longer. Another possibility would be to change the parameters programmed into AVEDac for what it should look at and what it should find or classify as interesting.

The improvements that would come from human detection would be practice or using experts. The testers used in this study had very little experience in identification

however they had very high detection rates, indicating that this would improve as they analyzed more video.

Classification Both humans and AVEDac had low rates of classification accuracy when classes were broken down into species identifications. This indicated that the difficulty in telling the differences between similar morphological species of holothurians is high for both human and machine. This was confirmed when the classes were combined into four classifications: Amperima-Peniagone Complex, Peniagone, Elpidia, and *Scotoplanes globosa*. Both humans and machine improved in their accuracy rates. However, human classification was still low while AVEDac showed a greatly improved accuracy rate. This was mostly due to the fact that it classified events that it was unsure of as "unknown" when the human testers were not given that option. If testers had been given this option, their accuracy rate may have improved but the percentage of correct would not have.

To improve these results in the future, the use of experienced test subjects would improve the data collected, especially if their experience is within holothurian identification. In respect to AVEDac's ability to classify, the size of the training library used in classification may improve the classification. TL Holothurian #5 had the smallest training libraries available for AVEDac to use for comparisons and the least accuracy in classification. The accuracy improved when libraries were combined to become larger. Though it is shown that library size will improve classification, a specific size that is needed for perfect identification is unknown and still needs to be researched. Future research would still benefit from larger libraries than those that were used in this study.

Overall, detection and classification can be improved within the AVEDac program by researching and testing which factors fully effect these two actions. For detection, which factors weight more within AVEDac: color, intensity, or orientation? And for classification, which factors determine class: motion, color, or texture? These things need to be determined and quantified in future research to improve the abilities of AVEDac.

CONCLUSIONS/RECOMMENDATIONS

In conclusion, we found that there are both advantages and disadvantages to both identification methods (human or machine). Humans need improvement in their classification accuracy, which can only come from practice and taking their time in classification, but their detection rates of organisms within videos is very high and accurate. AVEDac needs improvement in both the ability to detect and the classification of events and organisms. As stated before, the processing time and the parameters for determining what is interesting and what is not need to be quantified and improved.

AVEDac is a valuable tool and, though the detection and classification rates are not perfect and still need to be improved, the application of this program in research is still useful. AVEDac can cut down time needed to identify important events and creates a useful database for reference and comparison for researchers.

The only way to improve the machine classification process is to improve the human identification process. These two processes go hand and hand. This is why identifying the challenges with human identification is important to move forward in machine identification.

In addition to the improvements that were mentioned in the discussion, different classification methods should be tested to see the differences and which work better in the process. One of these methods includes the use of group testing instead of individual testing. This will open up discussion in how to identify organisms and events and improve the common understanding of what the organism or event is. Also, testing the training libraries within AVEDac in different color classifications will show differences in the classification by identifying the importance of color or lack thereof. Future testing in this question is needed because this experiment was only a preliminary test and can be greatly expanded upon for greater results and more in-depth questions.

ACKNOWLEDGEMENTS

I would like to thank Danelle Cline, my mentor, for her guidance throughout this project. She has helped me understand a very different aspect of research that I didn't

know about and has made my experience here at MBARI a fantastic one. She was always there when I had a question and was very easy to talk to. Also, she was very good at explaining things to me multiple times when I was confused.

I would also like to thank Linda Kuhnz for her help with the videos and explanations of species and how to identify things properly. I would have been lost without her help with the holothurian identification and classification. I would also like to thank her for her help in coordinating this internship and always helping us interns when we needed it.

Also, John Ryan for letting me use his LISST-HOLO data in my plankton analysis. I'm glad I got to help in the analysis of those images and that he allowed me to be one of the first to analyze them.

Most importantly, other than Danelle Cline, I'd like to thank George Matsumoto for everything that he did for me and for the other interns. He was a great guide to Monterey, telling me where I can get the best Boba tea and the best Italian food. As well as setting up the intern trips and the behind the scenes tours at the Aquarium, especially when we went to see the octopus tank. He made this experience an amazing one and something I will not forget.

I'd like to thank my fellow interns. I'd like to thank them for providing some companionship throughout this whole process and experience. I had a great time spending time with them and getting to know all of them. I wish them luck in everything that they do in the future.

Lastly, I must thank MBARI for giving me this opportunity to come out to the West coast and get hands on experience in this research field. Being here has renewed my excitement for the future and helped me see what I really want to do in life. It is incredible what this place can do in only ten weeks.

References:

Culvehouse, P.F., Gonzalez-Gil, S., Herry, V., Reguera, B., and Williams, R. (2003) Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. *Marine Ecology Progress Series*, 247: 17-25.

Culverhouse, P.F. (2007) Human and machine factors in algae monitoring performance. *Ecological Information*, 2: 361-366.

Kudela Lab. "Plankton Identification." *Plankton List*. CeNCOOS & HABMAP, n.d. Web. 30 Jun 2012.

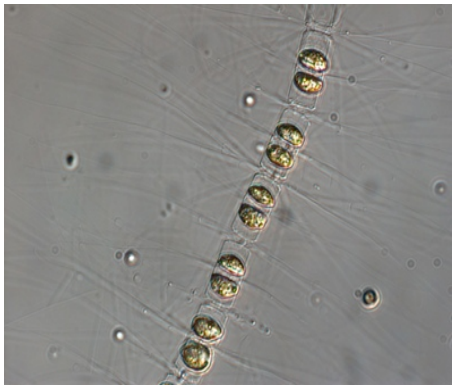
MBARI, (2012) Deep-Sea Guide. <seaspray.shore.mbari.org:8080/dsg/images/concept/Holothuroidea>. Web access: June-August 2012.

Perry, R. (2003) A Guide to the Marine Plankton of Southern California, 3rd Edition. *UCLA OceanGLOBE & Malibu High School*.

Schlining, B.M. and Stout, N.J. (2006) MBARI'S Video Annotation and Reference System. *Oceans*, 1-5.

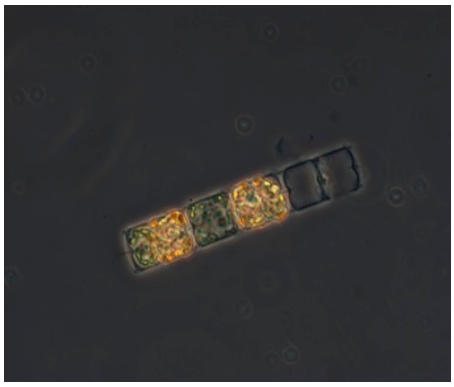
APPENDIX 1.

Diatoms



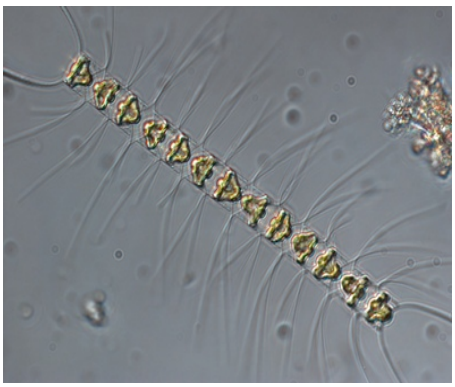
Bacteriastrum

Cylindrical cells bound together by the fusion of numerous setae that are regularly arranged around the cell margin. Cells have numerous small round chloroplasts.



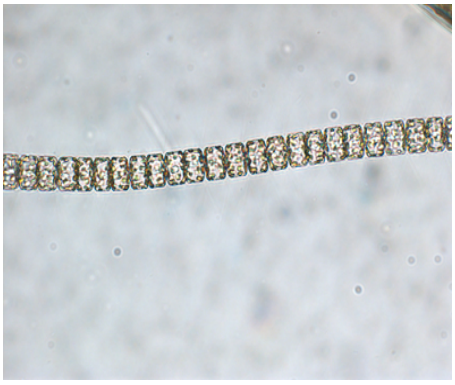
Cerataulina

Cells form close-set chains by apposition of tips of elevations. Cells have bipolar symmetry and bipolar elevations, and numerous small disk-like chloroplasts.



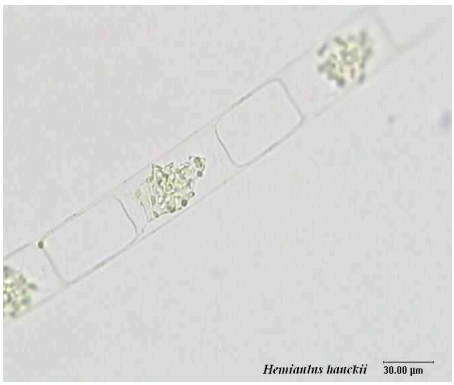
Chaetoceros

Cells form chains that are coiled, curved or straight. Long setae emerge from corners of cells.



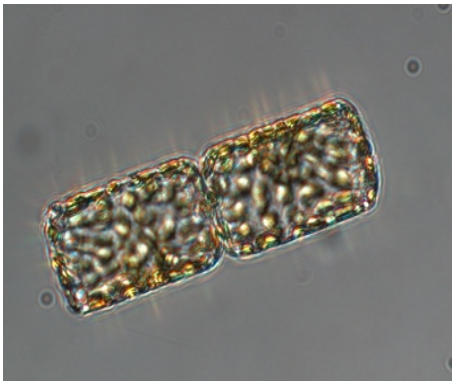
Detonula

Cells join together in mainly straight, stiff chains by short processes and mucilage threads.



Hemiaulus

Cells form chains, sometimes curved or twisted. Valves are elliptical with long and sometimes claw-like elevations (connecting horns).

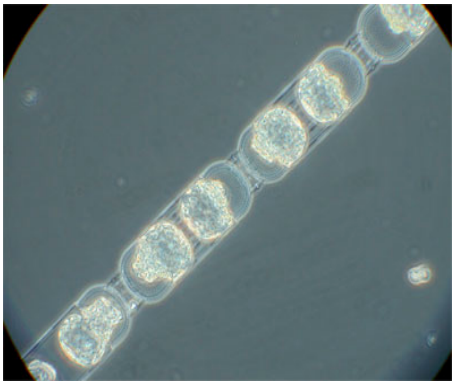


Lauderia

Centric diatom. Cells form chains, and are separated by occluded processes on marginal part of valve.

Skeletonema

Cells form chains, attached by external tubes or strutted processes organized in one marginal ring.



Stephanopyxis

Cylindrical, sometimes nearly spherical, capsule-shaped cells. Valves are domed with large hexagonal areolae.

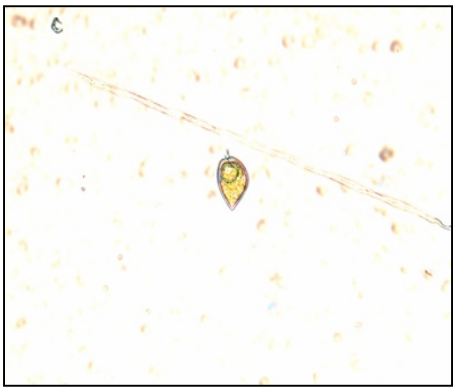
Thalassiosira

Chains-forming species are connected by a central organic thread. Numerous spine-like threads also extruded from strutted processes on valve margins.

Dinoflagellates

Noctiluca

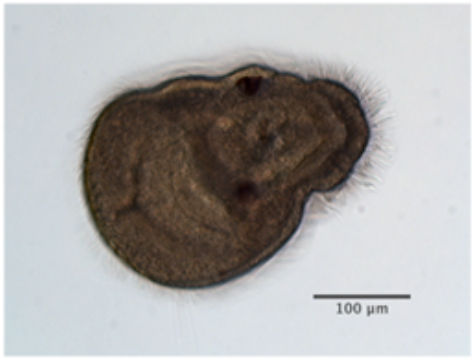
Large unarmored, round or kidney shaped cells with a striated tentacle, one flagellum and a eukaryotic nucleus.



Prorocentrum

Laterally compressed armored cell, ranging from heart-shaped to pyriform. There is no cingulum or sulcus. Well-developed apical spine.

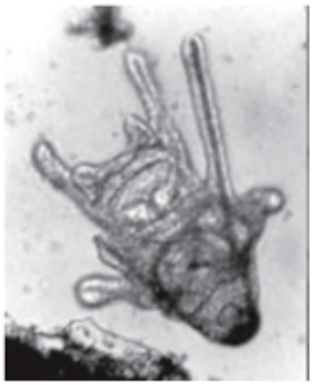
Bryozoa

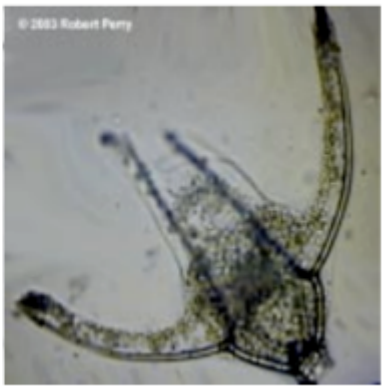


Copepod

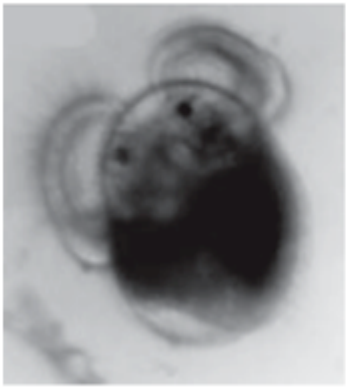
Crustacean

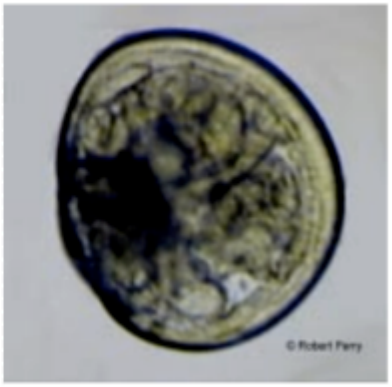
Echinodermata



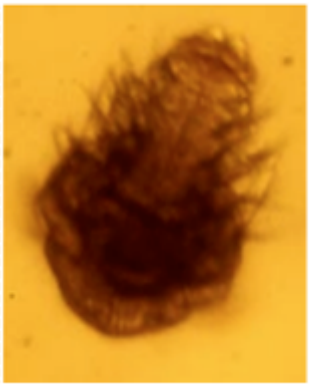


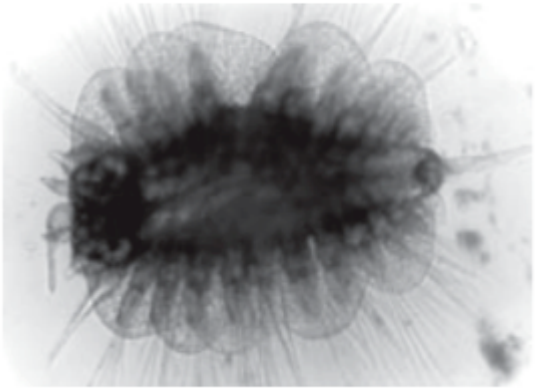
Mollusca



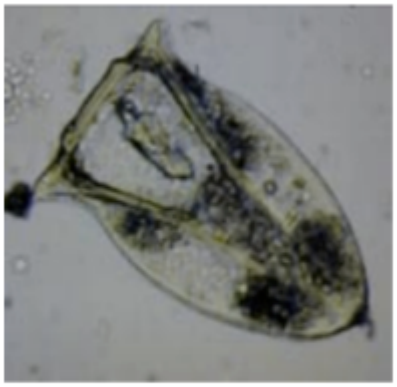


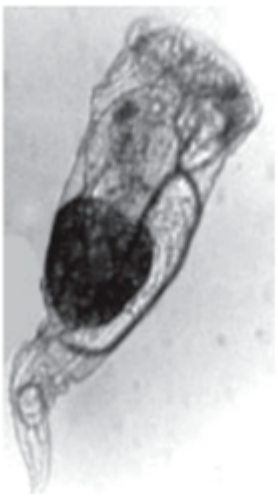
Polychaeta





Rotifera





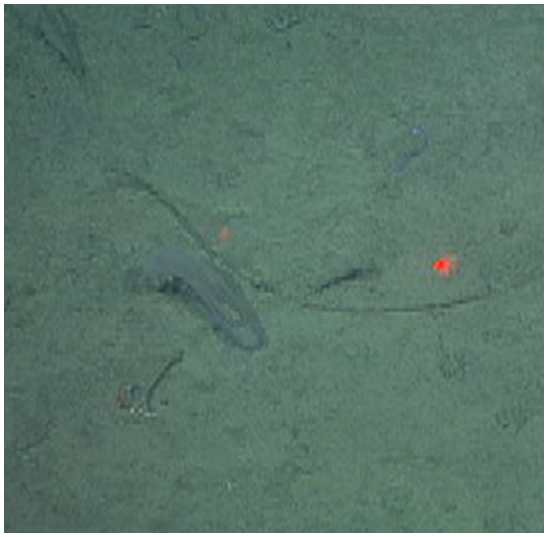
APPENDIX 2.

<< *Amperima-Peniagone* Complex

Orange-pink and somewhat opaque; turgid; feet allow body to stand above the seafloor; robust feeding tentacles; tips more darkly colored than body; two short dorsal fused parapopia at a sharp, nearly 90 degree inflection at the "head"

Peniagone sp. A >>

Transparent pink; flaccid; two long and two short dorsal fused parapodia; no sharp 90 degree inflection at "head"; robust feeding tentacles

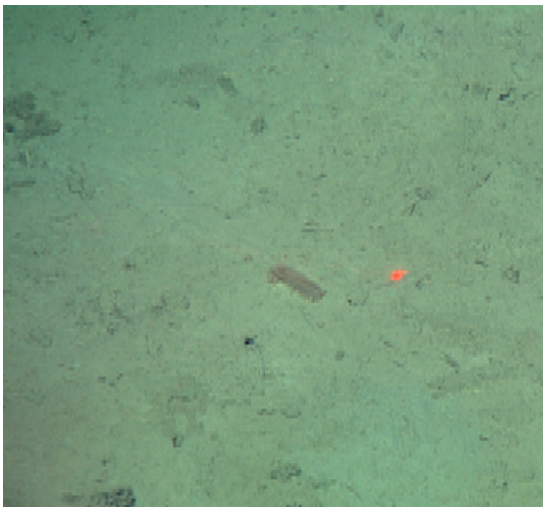


<< Peniagone sp. B

Transparent; flaccid; two short dorsal fused parapodia; no sharp 90 degree inflection at "head"; small, hard to see feeding tentacles

Peniagone sp. C >>

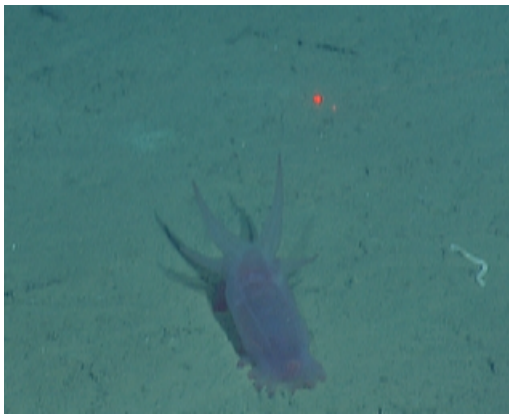
Transparent, sometimes tinged with purple/pink, long lateral parapodia, note very dark tips on feeding tentacles



<< Peniagone sp. D

Peniagone sp. E >>

Transparent pink, flaccid, very long fused dorsal parapodia



<< Elpidia

Light brown or sediment-colored; very small; stiff, projecting papillae

Scotoplanes globosa >>

APPENDIX 3.

Plankton:

1) What was one of the top deciding factors in the identification process (texture, shape, orientation, ect)?

2) On a scale of 1-10, 1being the lowest, rate the level of difficulty you feel it was to identify these images.

3) Did the comparison resource help you in the identification or would it have been more beneficial to have more images and/or images that were similar to the ones being tested?

4) Do you feel color would have changed/ influenced your predictions?

Holothurians:

1) What variables were used in identifying the holothurians when watching the video (color, texture, size, ect) and what were their order of importance to you?

2) When identifying species, what was the top variable you used to distinguish species?

3) Was the resource and descriptions helpful?

4) How difficult was it to identify holothurians within the video (scale of 1-10)?

5) How difficult was it to identify species within AVED (scale of 1-10)?

Both:

1) Did the quality of video/images aid in the ease/difficulty of identification?

2) Was it easier to identify video or still images?