Genetic Programming for Image Analysis and Upgrading Performance With the use of New Training Techniques

Hughes, James

***Abstract*—The focus of this work is to improve upon results obtained on image analysis done in assignment 2. Most system specific information was kept as consistent as possible between this work and the previous work. Image analysis was done on two different images. In one image the objective was to find boats and in the other image the objective was to find fossils. For both of these images the genetic programming system was given two images, a base image (the original image) and the other was a ground truth image (a key image used for validation). Multiple training techniques will be used in this work, all of which will use small sample areas obtained from the whole image. Testing will be done on the whole image from which the testing set came from. Many different experiments were done to determine which set of system parameters produces the best results. The fitness update technique that was used in this work was strictly Poli's fitness function on the early experimentation and a basic weighted sum was used for the larger experiments used for a concrete comparison of results between this and the previous work. In the previous work Poli's fitness function was only implemented on one of the images being analysed. The best performing training technique was training technique three and it did appear to produce better results than the experiments done in previous work but due to too many variables introduced it is difficult to confidently say that this work produced better results. Because of this, another set of experiments were done that used the same fitness parameters used for the larger set of comparisons done in the previous work. These two sets of experiments were then compared and it was found that one of the new training techniques did produce a definite improvement in fitness.**

I. Introduction

The purpose of this work is to continue exploring the ability of a genetic programming system (GP) to learn to find objects within an image. This work is a continuation of previous work done on image analysis [7]. This previous works objective was to have a GP system analyse an image to determine where certain objects exist and where they do not exist within the image. This work will be a continuation of the previous work in hopes to improve upon the results. In this work a dynamic fitness update strategy will be used for all experiments (in the previous work, it was only implemented for some of the experiments). This work will also focus on a new technique of training different than that of the previous work. These training techniques focus on isolating 'problem' areas within the image and trying to train more on areas that are more difficult for the system to understand. Similar to the previous work, many experiments will be conducted to determine which set of system parameters produces the best results. Ultimately the results in this work will be compared to the results obtained in the previous work to determine if this new training technique can yield better results. Because the results from the two sets of works will be compared to one another many system settings need to be left unchanged such as the language used by the GP system. The two images that were used for image analysis in the previous work will be the same images analysed in this work.

Multiple runs will be conducted for each set of system parameters to help produce statistical significance. Due to time constraints only ten sets of such experiments will be done on the first set of experiments as opposed to the thirty that are normally required. On the larger experimental analysis that will be used for a concrete comparison thirty runs will be done to ensure statistical significance.

Section II contains a definition of the problem sets used for the image analysis to be executed on the GP system. An overview of the fitness update strategy will be covered in section III along with an explanation of the variety of the new training techniques that will be used in the experimental analysis and the process in which the GP system learns. Section IV covers the methods of analysis that will be used in this work and section V covers the experimental results along with an analysis of their outcomes. Lastly section VI draws conclusions that can be derived from the experimental analysis and possible future works.

II. Problem Sets

Just like the previous works [7] two different images will be used for the experiments. The first image will be of boats in and around a marina in Miami Florida. The second image to be analysed will be of fossilized crinoids within rocks. The image of the marina was chosen because it is expected that the GP system will have a relatively easy time finding the objects within it. In the previous work the crinoid image was chosen because it was expected to be another easy image for the GP system to analyse due to the textures of the fossils within the image. Not only was it expected to be easy, but it also has some useful applications. Experimental analysis done in [7] showed that the GP system had a very difficult time trying to analyse the image but because a comparison of the results between this and the previous work will be done, the same images must be used. The results obtained in the previous work on this image were very bad so it is expected that the results produced in this work will be an easy improvement.

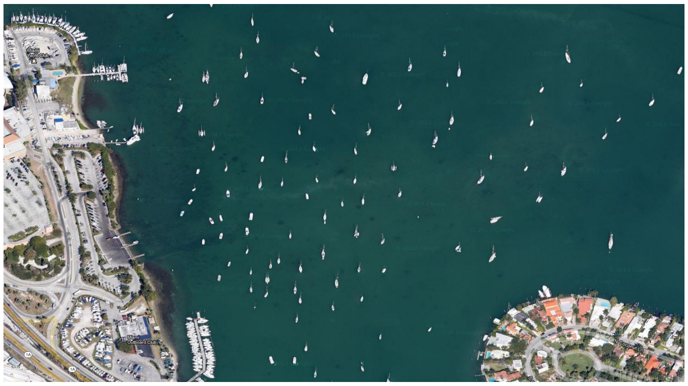


Fig. 1. Miami Image

A. Miami Image

The objects that are to be found for this image are boats in and around a marina in Miami Florida. The picture was originally obtained from Google Maps but the image was provided by Dr. Brian Ross and obtained for this work from the Genetic Programming course website [10]. Figure 1 shows the Miami image.

B. Fossil Image

The objects that are to be found in the fossil image are fossilized crinoids within rock. The picture was obtained from the petrified wood museum [12]. Refer to figure 2 for the image to be analysed. There are many different crinoids and they range in size. In figure 2 the crinoids are the circular objects and the objects that look similar to sand dollars. Notice how there is a variety of sizes of the crinoids in this image.

C. Applicability of Learning Strategies

Although there are other ideal machine learning algorithms for image analysis (neural networks), it has been shown that GP systems can also produce significant results[1] [9] . What makes GP systems interesting is that actual image analysis algorithms are produced by the system as opposed to other strategies that just are directly applied to an image.

III. Learning Algorithm

The basic ideas behind genetic algorithms and genetic programming will not be explained in detail in this work. For a detailed description refer to previous works [7]. The genetic programming system being used in this work is named Linkable GP [8] created by Michel Medland at Brock University.

A. Methods of Learning

The method of learning for a GP system for image analysis is simple. The first step is to obtain an image with objects that one wants to have the system try to learn to identify and an image that has the objects to be identified already found. This image with already identified objects is referred to the ground truth image and is used for the training and testing process.

Fig. 2. Fossilised Crinoid Image

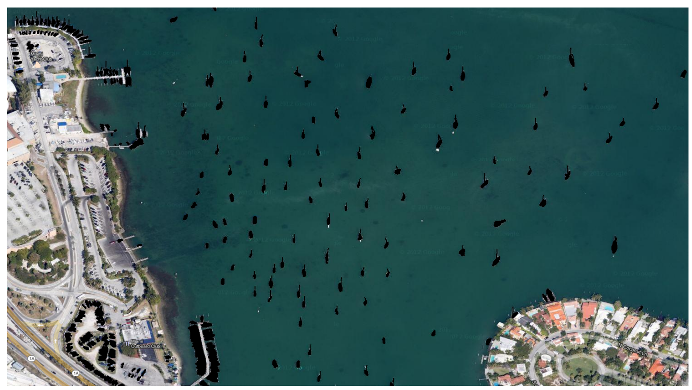


Fig. 3. Ground Truth Image for Boat Image

Refer to figures 3 and 4 for the ground truth images to be used in this work. In this work many small images are taken from the base image to produce the training set. The language that is used for image analysis can contain Boolean operations and mathematical equations. The most important functions for the language are ones that take a range of pixels and do some calculation based on the pixels characteristics.

The language used in this work will contain basic Boolean operations such as conditionals, greater than, less than, and & or. Mathematical operations will also be included such as addition, subtraction, multiplication, and division (the division ensures no divide by zero errors).

The terminals used in this work will be ephemeral constant doubles and boolean values (true and false).

The functions that use a range of pixels will each return a double value and they are as follows (Note, this is the same language used in assignment 2 [7]):

1) Average Hue: Determines the hue of each pixel in the provided range and returns the average hue value over the range which basically helps determine how close the colours in the range are to one another.

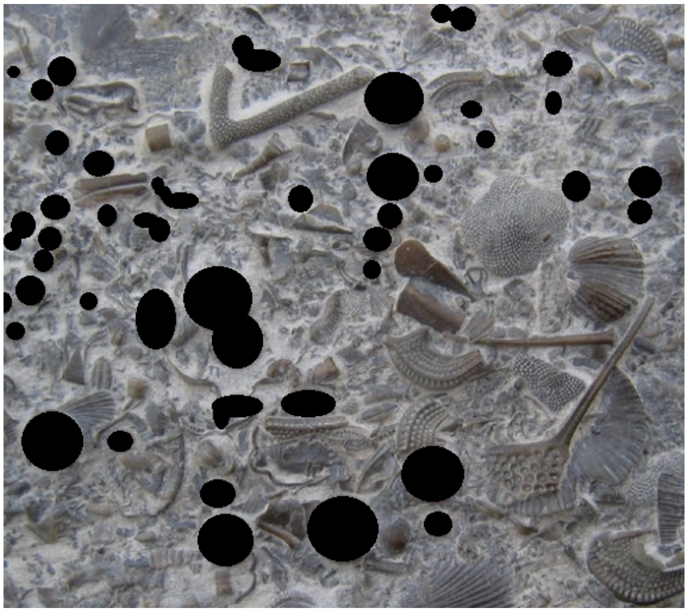


Fig. 4. Ground Truth Image for Crinoid Image (Note, poorly outlined, but needs to be the same as assignment 2)

2) Average Grey-scale: Calculates the grey-scale of each pixel in the provided range and returns the average grey-scale for the whole range.

3) Standard Deviation of Grey-Scale: Calculates the standard deviation of grey-scale within the provided range and returns this value. This operation uses the average grey-scale operation to determine the mean value for the standard deviation formula.

4) Average Brightness: Determines the brightness of each pixel and returns the average brightness of all pixels in the range.

5) Average Saturation: Determines the saturation (how much the light is distributed across the spectrum) for each pixel and returns the average saturation across the range provided.

6) Average Red, Green, Blue: Calculates the average red, green, blue (respective of which specific operation is called) in the pixels provided in the range and returns this value.

7) Average NTSC Grey-Scale: Exactly the same as the previously mentioned grey-scale operation but in this one it calculates the NTSC grey-scale as opposed to the basic grey-scale.

8) Standard Deviation of NTSC Grey-Scale: Exactly the same as the previously mentioned grey-scale standard deviation but in this one it calculates the standard deviation of the NTSC grey-scale.

For a full printout of the languages used for this work refer to Appendix II.

The learning process occurs when the language tree makes a prediction and we determine if its prediction is correct by referring to the ground truth image. This process takes place over the whole test image (many pixels) and the fitness of the language tree is determined through some fitness update strategy. The selection process and genetic operations occur here but their explanation has been omitted, refer to previous works for a detailed explanation [6].

B. Fitness Update Strategies

1) Poli's Fitness Function: The fitness update strategy used in this work is Poli's fitness update strategy [9]. The function does not have an official name but it is being referred to as Poli's fitness in this work for easy referral. This fitness update strategy was used in the previous work with success on the boat image analysis but was not used on the fossil image analysis due to poor performance. In this work Poli's fitness update strategy will be used on the fossil image analysis as preliminary testing showed that with some adjustments the strategy could be used with more success. This fitness function dynamically adjusts what is considered 'good' results which helps the system actively fine tune the learning process. The formula is as follows:

$$f(x) = FP + FN * exp^{(10(\frac{FN}{P} - \alpha))} \tag{1}$$

$$Where$$

$$FP\ is\ the\ number\ of\ false\ positives$$

$$FN\ is\ the\ number\ of\ false\ negatives$$

$$p\ is\ the\ number\ of\ positive\ examples$$

$$\alpha\ is\ the\ domain\ dependent\ parameter$$

$$which\ represents\ the\ maximum\ percentage\ of$$

$$misses\ above\ which\ the\ number\ of\ misses$$

$$is\ considered\ unacceptable$$

2) Weighted Sum: A weighted sum is arguably the simplest form of fitness update for multi-objective optimization. The formula for the weighted sum is as follows:

$$WeightedSum = \sum_{i=1}^{n} w_i a_i \tag{2}$$

$$Where$$

$$n\ is\ the\ number\ of\ dimensions$$

$$w_i\ is\ the\ weight\ corresponding\ to\ the\ i^{th}\ value$$

$$a_i\ is\ thenumber\ of\ occurences\ of\ the\ i^{th}\ value$$

This method applies a weight to objectives and then adds the values together. The use of weights is to help the system compensate for objectives that may overshadow others. In this work, there are a lot more pixels that belong to not the objects to be found so the system would simply classify everything as not the object to be found and it would receive a good fitness value. By adding a weight for correctly classified pixels the fitness function will give a more 'fare' fitness value that represents the accuracy of classification better.

C. Training Techniques

Three separate training techniques will be used in this work and each of which will have a set of five small training images that are 80x80 (6400 total) pixels each. These five small images were chosen based on the content within them and how well the algorithm had previously categorized the objects within that area of the image (i.e. known problem areas were chosen).

1) Technique 1: For the first technique, each chromosome would only see one of these five images (selected with a random shuffle) (can train on a different training image in the next generation) and their fitness would be compared to one another when the selection process occurs. The strategy behind this training technique is to reduce learning time through training on many small images and to improve results by allowing the system to learn on known problem areas.

2) Technique 2: The next training technique is very similar to the first but with one major difference. When the run has completed half of its generations the two training images that the GP system is performing the best on are removed from the training area. The GP system keeps a record of which image had the lowest average fitness (lower fitness is good as it is error in this work) per generation. When the two best images are removed from the training set this bookkeeping still occurs so when another image has more generations where it had the lowest average fitness it will then be removed from the training set and the one it passed will re-enter. The motivation for using this technique is similar to the first technique but also because it will prevent the system from over-training on certain training areas and allow more training on more difficult areas.

3) Technique 3: The last of the training techniques to be used in this work is similar to the previous two training techniques. Instead of a chromosome only seeing on one of the five images per generation, a chromosome will see all five images per generation. Similar to the last technique, once the run has completed half of its generations the current two training images that the GP system performs the best on are removed from the training set. Once again, this is done with bookkeeping. The reason for implementing this technique is to see how a GP system will do when allowed to train on more pixels per chromosome compared to one that only trains on one 80x80 pixel image. The removal of the two current best training images is done to prevent over-training and allow more training on problem areas. Another important reason for the removal is because not only does GP learning take a long time [3], as a run goes on bloat normally occurs which dramatically increases run times [4]. Removing two of the five training images from the whole training pool will reduce training time which ultimately will reduce GP run times.

IV. Methods of Analysis

The goal of this work is to improve upon the results obtained in assignment 2 (cite ass2). Before a comparison is done between the two works the best results need to be found in this work (which training technique works best). The statistical tests that will be conducted in this work will be mean analysis, standard deviation, a 90% confidence interval, ANOVA, and a paired t-test (Mann-Whitney U test may be required depending on normal distribution). Refer to previous works for an explanation of these statistical tests [6] [7].

V. Results and Discussion

This section contains an analysis of all experiments done over the new training techniques on both the boat image and the crinoid image. Subsections V-A, V-B, and V-C covers an analysis of each training technique on both images. Subsection V-D does a comparison between the techniques and determines which performed the best. Subsection V-E does a comparison between the best results obtained in this work and the best results obtained in assignment 2. Lastly, subsection V-F will contain the experimental results for the crinoid image analysis using fitness sharing and the same system parameters and settings as assignment 2 and a comparison between the two assignment's results. Only the noteworthy experiments will be considered in this section.

The set of five small training images for both the boat and crinoid images can be seen in Appendix I. Each of these images are 80x80 pixels (6400 pixels). This means for the first two training techniques each chromosome will only see 6400 pixels per generation. For training technique 3 the chromosomes will see 32000 pixels per generation as in this technique the chromosomes see all five images. The testing sets are 1184641 pixels and 290004 pixels for the boat and crinoid images respectively.

For all early experiments the learning curves were left out as they were not descriptive of the results. Due to the nature of Poli's fitness update strategy the average results were normally infinity and because elitism was not implemented so the graphs were chaotic. The last set of experiments using the weighted sum fitness update strategy will have the learning curves included.

The last set of experiments to be done with the weighted sum fitness update technique will only be done on the crinoid image. The reason for this is because the results obtained for the boat image analysis in assignment 2 were reasonably good and the results for the crinoid image in assignment 2 were very bad. Because the results for the crinoid image were so bad, it would be ideal to try to focus on this image and improve its results.

Just like in assignment 2, the fitness is what is being optimized, not the direct number of correct classifications which is ultimately what we want. Sometimes the best classification results do not directly correspond to the best fitness value.

Unlike assignment 2, only the base image for the crinoid image analysis will be used. In assignment 2, the images with the filters (edge detection and sharpness) did not produce better results.

For all experiments a tournament selection of size three was used. All languages in this paper will be simplified versions as the originals are enormous.

Refer to Appendix I for all of the best images obtained for each of the best performing training sets.

-	Gens	Pop	Cross	Mut	Min	Max	Radius	α
C1	20	50	.4	.7	5	10	1	.7
C2	20	50	.4	.7	5	10	2	.7
C3	20	50	.5	.8	5	7	1	.7
C4	20	50	.5	.8	5	7	2	.7
C5	20	50	.7	.7	5	10	1	.7
C6	20	50	.7	.7	5	10	2	.7
C7	20	50	.7	.9	5	10	1	.7
C8	20	50	.75	.5	5	10	1	.7
C9	20	50	.75	.5	5	10	1	.65
C10	20	50	.75	.5	5	10	2	.7
C11	20	50	.75	.5	5	10	2	.65
C12	20	50	.75	.5	5	10	2	.75

TABLE I
SYSTEM PARAMETERS FOR BOAT IMAGE ANALYSIS WITH TRAINING TECHNIQUE ONE (NOTE: MIN AND MAX REFER TO TREE DEPTH)

```
One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12

Source   DF        SS      MS     F      P
Factor   11   1448654  131696  0.44  0.933
Error   108  32052608  296783
Total   119  33501262

S = 544.8   R-Sq = 4.32%   R-Sq(adj) = 0.00%

                              Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean   StDev     +---------+---------+---------+---------
C1     10  5970.6   743.5              (----------*----------)
C2     10  6073.1   231.2                  (----------*-----------)
C3     10  5975.1   315.8              (----------*-----------)
C4     10  6034.6   290.7                 (-----------*-----------)
C5     10  5880.9   811.3          (-----------*----------)
C6     10  6004.9   278.0                (-----------*-----------)
C7     10  6059.7   532.1                  (----------*-----------)
C8     10  6009.4   276.6                (-----------*-----------)
C9     10  6152.9   199.8                       (----------*-----------)
C10    10  5729.3  1217.6     (----------*----------)
C11    10  5981.7   309.7              (-----------*-----------)
C12    10  6148.4   240.1                       (----------*-----------)
                              +---------+---------+---------+---------
                           5400      5700      6000      6300

Pooled StDev = 544.8
```

Fig. 5. ANOVA results for training technique 1 on the Miami image.

A. Training Technique One

1) Miami Image: For the first training technique done on the boat image there were 12 separate experiments done, each with a different set of system parameters. Refer to table I for the system parameters used for each of the experiments. An ANOVA test was done on these experiments to help rank the results. See figure 5 for the ANOVA results. Experiments C1-C12 correspond to the experiments listed in table (ref table again).

When analysing the ANOVA results seen in figure 5 we can see that all results were relatively close to one another but experiment C10 was the best performing based on a basic mean analysis and C5 performed the second best. When doing a paired t-test and a Mann-Whitney U test it is observed that there is no statistical difference between these two results (there is actually no statistical difference between any of the results).

Figure 6 shows a graphical summary of experiment C10. This graphical summary covers all of the important information regarding the experiment. Once again, the learning curve will not be included. For the purpose of future comparisons only experiment C10 will be considered (it has the best mean value) even though there are no differences between the obtained results.

When referring to figure 6 one can that the mean value

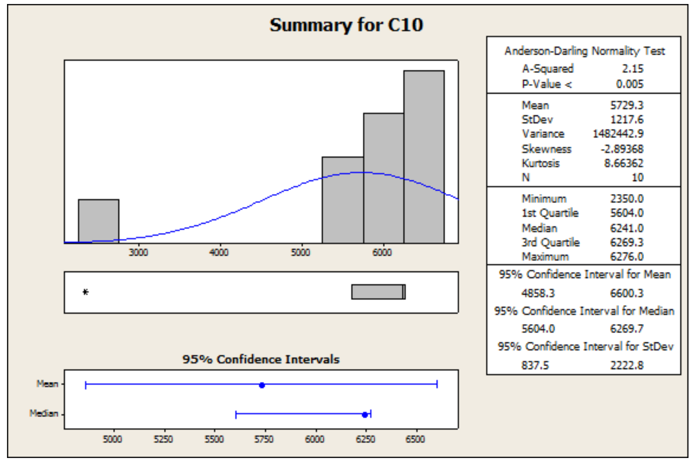


Fig. 6. Graphical Summary for Experiment C10 Which had the Best Results

-	True	False
Positive	26538	127884
Negative	1080444	4489

TABLE II
% CORRECT POS: 0.8553, % CORRECT NEG: 0.8941, OVERALL % ACCURACY: 0.8931

obtained was 5729.3 and the median was 6241. The best obtained was 2350. The standard deviation was very high at 1217.6. The confidence interval range is also very high and there is no clear normal distribution evident.

Refer to Appendix I for the best result image obtained for experiment C10. Just like in assignment 2 there are some trends that can be noted, namely the edge cases. When ignoring edge cases the system found all of the boats rather effectively but the GP system got very confused by the land.

The GP system produced the following language for the best result in experiment C10 (simplified):

```
(avgRed > (306.824 - avgHue - stdDev)
```

The confusion matrix for the best result in experiment C10 can be seen in table II.

2) Crinoid Image: Just like the previous experiments, many individual experiments were done, each with different system parameters. There were 12 such experiments and table III shows the system parameters for each experiment.

Refer to figure 7 for the ANOVA output for all of these experiments. Just like the previous experiment all of the results are very similar and doing more statistical analysis shows that there is no statistical difference in their results. When doing a basic mean analysis we can see that experiment C5 had the best results and we will consider this experiment for comparisons between the different training techniques.

Figure 8 is the graphical summary for experiment C5 for the first training technique on the crinoid image. The mean value obtained was 2384.9 and the median is very close to the mean at 2298. The standard deviation and confidence intervals are lower than that of the boat image analysis but the results show that the experiments are just consistently bad together.

The language obtained in the best performing experiment

-	Gens	Pop	Cross	Mut	Min	Max	Radius	α
C1	20	50	.4	.7	5	10	1	.7
C2	20	50	.4	.7	5	10	2	.7
C3	20	50	.5	.8	5	7	1	.7
C4	20	50	.5	.8	5	7	2	.7
C5	20	50	.7	.7	5	10	1	.7
C6	20	50	.7	.7	5	10	2	.7
C7	20	50	.7	.9	5	10	1	.7
C8	20	50	.75	.5	5	10	1	.65
C9	20	50	.75	.5	5	10	1	.75
C10	20	50	.75	.5	5	10	2	.7
C11	20	50	.75	.5	5	10	2	.65
C12	20	50	.75	.5	5	10	2	.75

TABLE III
SYSTEM PARAMETERS FOR CRINOID IMAGE ANALYSIS WITH TRAINING TECHNIQUE ONE (NOTE: MIN AND MAX REFER TO TREE DEPTH)

```
One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12

Source   DF        SS      MS     F      P
Factor   11   4181474  380134  1.17  0.318
Error   108  35145070  325417
Total   119  39326545

S = 570.5   R-Sq = 10.63%   R-Sq(adj) = 1.53%

                               Individual 95% CIs For Mean Based on
                               Pooled StDev
Level   N    Mean  StDev  --+---------+---------+---------+-------
C1     10  2485.8  586.8     (---------*---------)
C2     10  2673.5  374.7          (---------*----------)
C3     10  2810.0  582.5              (---------*----------)
C4     10  2962.6  706.9                  (----------*---------)
C5     10  2384.9  522.9  (---------*---------)
C6     10  2918.7  795.1                 (---------*----------)
C7     10  2496.5  154.9     (---------*----------)
C8     10  2599.3  634.7        (---------*---------)
C9     10  2404.4  310.0  (----------*---------)
C10    10  2513.8  565.5      (---------*---------)
C11    10  2717.4  618.2           (----------*---------)
C12    10  2774.8  671.3             (---------*---------)
                          --+---------+---------+---------+-------
                          2100      2450      2800      3150

Pooled StDev = 570.5
```

Fig. 7. ANOVA results for training technique 1 on the Crinoid image.

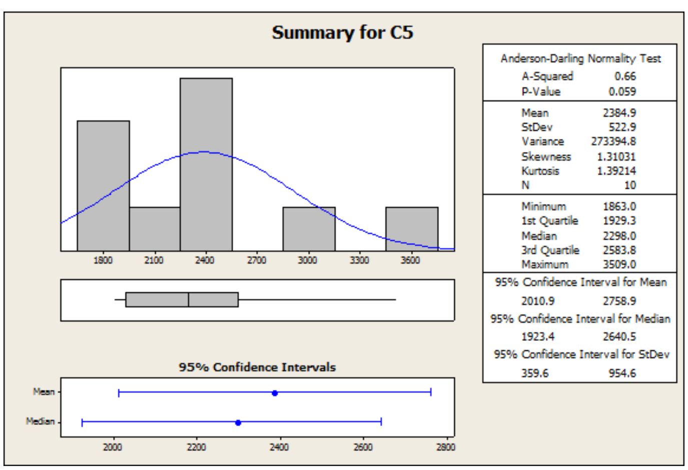


Fig. 8. Graphical Summary for Experiment C5 Which had the Best Results

-	True	False
Positive	21843	125715
Negative	133935	8511

TABLE IV
% CORRECT POS: 0.7196, % CORRECT NEG: 0.5158, OVERALL % ACCURACY: 0.5371

-	Gens	Pop	Cross	Mut	Min	Max	Radius	α
C1	20	50	.4	.7	5	10	1	.7
C2	20	50	.4	.7	5	10	5	.7
C3	20	50	.5	.8	5	7	1	.7
C4	20	50	.5	.8	5	7	5	.7
C5	20	50	.7	.7	5	10	1	.7
C6	20	50	.7	.7	5	10	5	.7
C7	20	50	.7	.9	5	10	1	.7
C8	20	50	.75	.9	5	10	5	.7
C9	20	50	.75	.5	5	10	1	.75
C10	20	50	.75	.5	5	10	1	.65
C11	20	50	.75	.5	5	10	2	.7
C12	20	50	.75	.5	5	10	2	.65
C13	20	50	.75	.5	5	10	2	.75

TABLE V
SYSTEM PARAMETERS FOR BOAT IMAGE ANALYSIS WITH TRAINING TECHNIQUE TWO (NOTE: MIN AND MAX REFER TO TREE DEPTH)

in the set C5 is as follows (simplified):

```
(avgBlue * avgSaturation) >
     (avgBlue - avgRed)
```

And the confusion matrix can be seen in table IV.

Training technique one did not seem to be very effective for both the boat and crinoid image analysis experiments.

B. Training Technique Two

1) Miami Image: Thirteen unique experiments were down for the second training technique done on the boat image. Each of these thirteen experiments has a different set of system parameters. Refer to table V for the system parameters for each of these thirteen experiments.

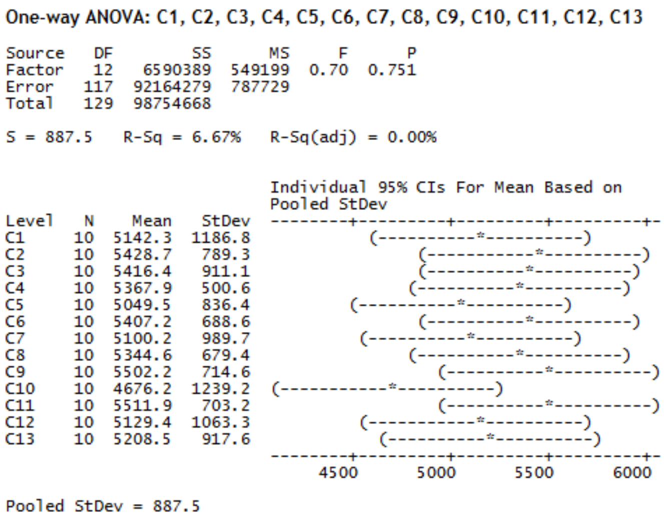
Figure 9 has the ANOVA output for this set of experiments. Experiment C10 has the best mean value of all of the experiments done but once again, all of the experiments seem very close to one another. A Paired t-test and Mann-Whitney U tests were done and they showed no statistical difference between the best performing results. Regardless of this lack of statistical differences between these experiments, experiment C10 will be used for a higher level comparison due to its mean value.

Figure 10 shows a graphical summary of experiment C10. There is no normal distribution in this experiment. The mean value obtained was 4676.2 and the median was 5168. The standard deviation and 90% confidence interval were rather large showing that these results are chaotic.

The language obtained in the best performing experiment in the set C10 is as follows (simplified):

```
(stDev * 12.63) >=
     ( 104.15)
```

The confusion matrix can be seen in table VI.


```
One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13

Source   DF        SS      MS     F      P
Factor   12   6590389  549199  0.70  0.751
Error   117  92164279  787729
Total   129  98754668

S = 887.5   R-Sq = 6.67%   R-Sq(adj) = 0.00%

                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level   N    Mean   StDev
C1     10  5142.3  1186.8
C2     10  5428.7   789.3
C3     10  5416.4   911.1
C4     10  5367.9   500.6
C5     10  5049.5   836.4
C6     10  5407.2   688.6
C7     10  5100.2   989.7
C8     10  5344.6   679.4
C9     10  5502.2   714.6
C10    10  4676.2  1239.2
C11    10  5511.9   703.2
C12    10  5129.4  1063.3
C13    10  5208.5   917.6
                                  4500      5000      5500      6000

Pooled StDev = 887.5
```

Fig. 9. ANOVA results for training technique 2 on the Miami image.

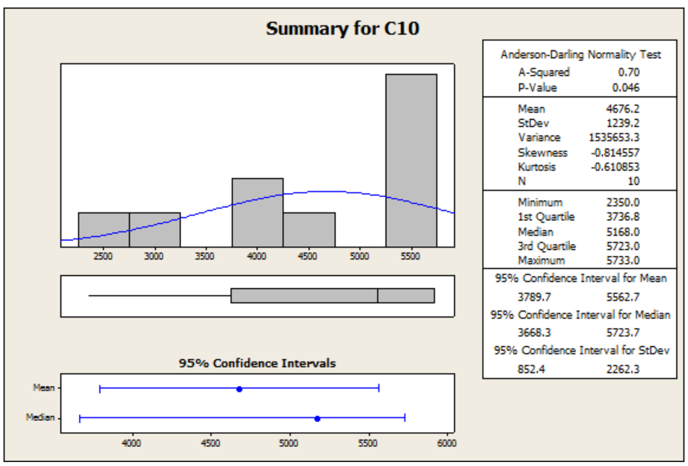
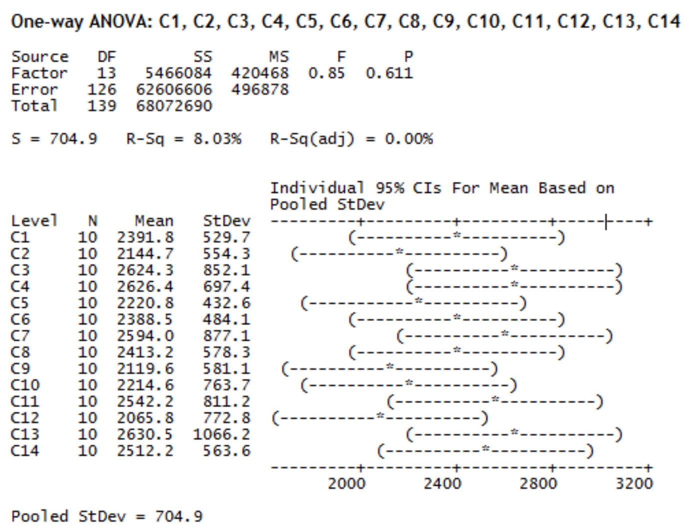


Fig. 10. Graphical Summary for Experiment C10 Which had the Best Results

-	Gens	Pop	Cross	Mut	Min	Max	Radius	α
C1	20	50	.4	.7	5	10	1	.7
C2	20	50	.4	.7	5	10	2	.7
C3	20	50	.4	.7	5	10	5	.7
C4	20	50	.5	.8	5	7	1	.7
C5	20	50	.5	.8	5	7	2	.7
C6	20	50	.5	.8	5	7	5	.7
C7	20	50	.7	.7	5	10	1	.7
C8	20	50	.7	.7	5	10	2	.7
C9	20	50	.7	.9	5	10	1	.7
C10	20	50	.75	.5	5	10	1	.65
C11	20	50	.75	.5	5	10	1	.75
C12	20	50	.75	.5	5	10	2	.7
C13	20	50	.75	.5	5	10	2	.65
C14	20	50	.75	.5	5	10	2	.75

TABLE VII
SYSTEM PARAMETERS FOR CRINOID IMAGE ANALYSIS WITH TRAINING TECHNIQUE TWO (NOTE: MIN AND MAX REFER TO TREE DEPTH)


```
One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13, C14

Source   DF        SS      MS     F      P
Factor   13   5466084  420468  0.85  0.611
Error   126  62606606  496878
Total   139  68072690

S = 704.9   R-Sq = 8.03%   R-Sq(adj) = 0.00%

                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level   N    Mean   StDev
C1     10  2391.8   529.7
C2     10  2144.7   554.3
C3     10  2624.3   852.1
C4     10  2626.4   697.4
C5     10  2220.8   432.6
C6     10  2388.5   484.1
C7     10  2594.0   877.1
C8     10  2413.2   578.3
C9     10  2119.6   581.1
C10    10  2214.6   763.7
C11    10  2542.2   811.2
C12    10  2065.8   772.8
C13    10  2630.5  1066.2
C14    10  2512.2   563.6
                                  2000      2400      2800      3200

Pooled StDev = 704.9
```

Fig. 11. ANOVA results for training technique 2 on the Crinoid image.

2) Crinoid Image: Fourteen unique experiments were done for this training technique on the crinoid image each with a unique set of system parameters. Table VII shows their system parameters.

The ANOVA results can be seen in figure 11. Again, these results are very similar and there does not seem to be much difference between any of these results. The best performing set of experiments is C12 and the second best is C10. Their paired t-test resulted in a p-value of 0.622 and the Mann-Whitney U test's p-value was 0.9698 which shows no difference between these experiments. Once again, even though there is no difference, experiment set C12 will be used for comparison between training techniques due to its mean value.

Figure 12 shows a graphical summary for experiment set C12. There is a normal distribution in this experiment set. The standard deviation and the 90% confidence interval is still large but lower than other experiments.

The language obtained in the best performing experiment in the set C12 is as follows (simplified):

```
((238.9785 - avgRed) > (avgBlue)) OR
     ((avgRed) > (avgHue/1.049))
```

The confusion matrix can be seen in figure VIII.

Just like training technique one, this one does not appear to have promising results for both the boat and crinoid

-	True	False
Positive	26353	168275
Negative	1040053	4674

TABLE VI
% CORRECT POS: 0.8493, % CORRECT NEG: 0.8607, OVERALL % ACCURACY: 0.8604

-	True	False
Positive	20225	143753
Negative	115897	10129

TABLE VIII
% CORRECT POS: 0.6663, % CORRECT NEG: 0.4463, OVERALL % ACCURACY: 0.4693

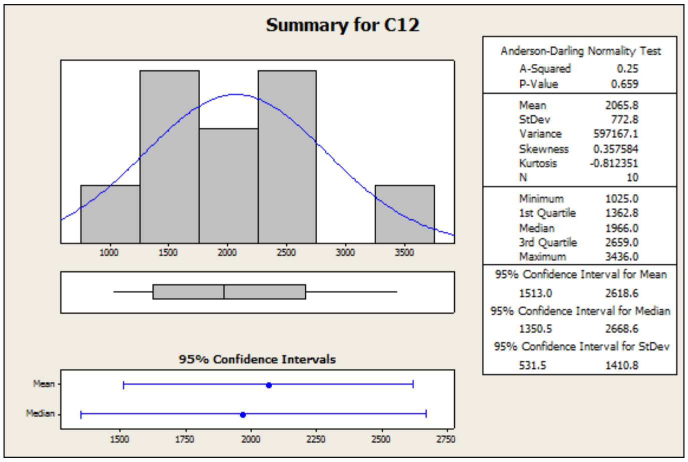


Fig. 12. Graphical Summary for Experiment C12 Which had the Best Results

-	Gens	Pop	Cross	Mut	Min	Max	Radius	α
C1	20	50	.4	.7	5	10	1	.7
C2	20	50	.4	.7	5	10	2	.7
C3	20	50	.4	.7	5	7	5	.7
C4	20	50	.5	.8	5	10	5	.7
C5	20	50	.7	.7	5	10	5	.7
C6	20	50	.7	.9	5	10	1	.7
C7	20	50	.7	.9	5	10	5	.7
C8	20	50	.75	.5	5	10	1	.65
C9	20	50	.75	.5	5	10	1	.75
C10	20	50	.75	.5	5	10	2	.7
C11	20	50	.75	.5	5	10	2	.65
C12	20	50	.75	.5	5	10	2	.75

TABLE IX
SYSTEM PARAMETERS FOR BOAT IMAGE ANALYSIS WITH TRAINING TECHNIQUE TWO (NOTE: MIN AND MAX REFER TO TREE DEPTH)

image analysis. This is not overly surprising due to the very small number of training pixels each chromosome sees per generation.

C. Training Technique Three

1) Miami Image: Twelve experiment sets were done for the third training technique on the boat image, each with a unique set of system parameters. Table IX shows the system parameters for each of these experiment sets.

The ANOVA results for this group of experiments can be seen in figure 13. Once again all experiments seem to have similar results. Experiment set C6 has the best mean value. The second best performing results were from experiment C1 and a paired t-test was done to determine if there is any statistical difference between these results. The p-value for the paired t-test was 0.073 and 0.0757 for a Mann-Whitney U test which is small, but not small enough to have a 95% confidence that there is a significant difference in their results. This paired t-test had the smallest p-value of all experiments so far.

Figure 14 shows a graphical summary of the results for the best performing experiment set C6. This experiment has a somewhat normal distribution and a relatively small standard deviation and 90% confidence interval (relative to other experiment sets). Once again, even though there is no statistical significant difference between the results obtained in

```
One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12

Source   DF         SS       MS     F      P
Factor   11   18535083  1685008  1.48  0.150
Error   108  123103789  1139850
Total   119  141638872

S = 1068   R-Sq = 13.09%   R-Sq(adj) = 4.23%

                            Individual 95% CIs For Mean Based on Pooled StDev
Level   N  Mean  StDev  --+---------+---------+---------+-------
C1     10  2862   1046           (--------*-------)
C2     10  3165    449                (--------*-------)
C3     10  3409    927                    (--------*-------)
C4     10  3970   1470                             (--------*-------)
C5     10  3083    768               (--------*-------)
C6     10  2195    587  (-------*--------)
C7     10  3193    750                  (-------*-------)
C8     10  3187   1484                (--------*-------)
C9     10  3268   1343                  (--------*-------)
C10    10  3093    769               (--------*-------)
C11    10  2837   1296           (-------*--------)
C12    10  3214   1261                  (-------*--------)
                        --+---------+---------+---------+-------
                        1600      2400      3200      4000

Pooled StDev = 1068
```

Fig. 13. ANOVA results for training technique 3 on the Miami image.

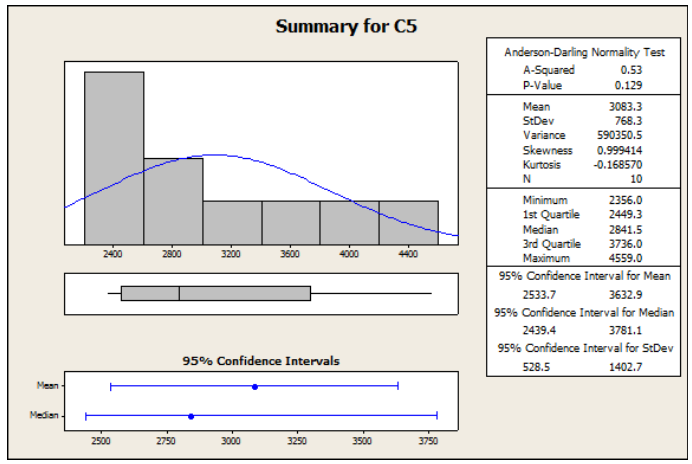


Fig. 14. Graphical Summary for Experiment C6 Which had the Best Results

this experiment set, experiment set C6 will be used for higher level comparisons due to the fact it does have the lowest mean value (3083.3).

The language obtained in the best performing experiment set in C6 is as follows (simplified):

```
if(avgGrey == stdDevNTSC)
     Then TRUE
     Else ((avgGreyNTSC + 72.5389) <
           (stdDevNTSC))
```

The confusion Matrix can be seen in table X.

2) Crinoid Image: This last group of experiments had a total of 12 unique individual experiments done, each with different system parameters. Table XI has a description of the system parameters for each group.

-	True	False
Positive	11944	31169
Negative	1177159	19083

TABLE X
% CORRECT POS: 0.3849, % CORRECT NEG: 0.9742, OVERALL % ACCURACY: 0.9594

-	Gens	Pop	Cross	Mut	Min	Max	Radius	α
C1	20	50	.4	.7	5	10	1	.7
C2	20	50	.4	.7	5	10	2	.7
C3	20	50	.4	.7	5	10	5	.7
C4	20	50	.5	.8	5	7	5	.7
C5	20	50	.7	.7	5	10	5	.7
C6	20	50	.7	.9	5	10	1	.7
C7	20	50	.7	.9	5	10	5	.7
C8	20	50	.75	.5	5	10	1	.65
C9	20	50	.75	.5	5	10	1	.75
C10	20	50	.75	.5	5	10	2	.7
C11	20	50	.75	.5	5	10	2	.65
C12	20	50	.75	.5	5	10	2	.75

TABLE XI
SYSTEM PARAMETERS FOR CRINOID IMAGE ANALYSIS WITH TRAINING TECHNIQUE TWO (NOTE: MIN AND MAX REFER TO TREE DEPTH)

One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13, C14

```
Source   DF        SS      MS     F      P
Factor   13   5466084  420468  0.85  0.611
Error   126  62606606  496878
Total   139  68072690

S = 704.9   R-Sq = 8.03%   R-Sq(adj) = 0.00%

                               Individual 95% CIs For Mean Based on
                               Pooled StDev
Level   N    Mean   StDev  ---------+---------+---------+---------+
C1     10  2391.8   529.7          (----------*----------)
C2     10  2144.7   554.3    (----------*----------)
C3     10  2624.3   852.1                (----------*----------)
C4     10  2626.4   697.4                (----------*----------)
C5     10  2220.8   432.6     (----------*----------)
C6     10  2388.5   484.1          (----------*----------)
C7     10  2594.0   877.1               (----------*----------)
C8     10  2413.2   578.3          (----------*----------)
C9     10  2119.6   581.1   (----------*----------)
C10    10  2214.6   763.7     (----------*----------)
C11    10  2542.2   811.2              (----------*----------)
C12    10  2065.8   772.8  (----------*----------)
C13    10  2630.5  1066.2                (----------*----------)
C14    10  2512.2   563.6             (----------*----------)
                           ---------+---------+---------+---------+
                                 2000      2400      2800      3200

Pooled StDev = 704.9
```

Fig. 15. ANOVA results for training technique 3 on the Crinoid image.

The ANOVA results for this group of experiments can be seen in figure 15. Once again, like every group of experiments beforehand, all experiments are close to each other in the fitness landscape. A paired t-test was done between the best performing experiment sets and it showed that there is no statistical significant difference between these results. That being said, there is a difference between the best and the worst performing experiment sets. Like all experiments before, only experiment C5 will be used for higher comparisons as it had the lowest mean value .

Figure 16 is the graphical summary of experiment set C5. There is a normal distribution and the standard deviation and 95% confidence interval are the lowest seen so far, but still large. The median value obtained was 1386.0 and the mean value was 1683.9.

The language obtained in the best performing experiment set in C6 is as follows(simplified):

```
(avgBlue * avgBrightness) > 62.9713)
```

The confusion matrix can be seen in table XII.

This training technique appears to produce better results than the others.

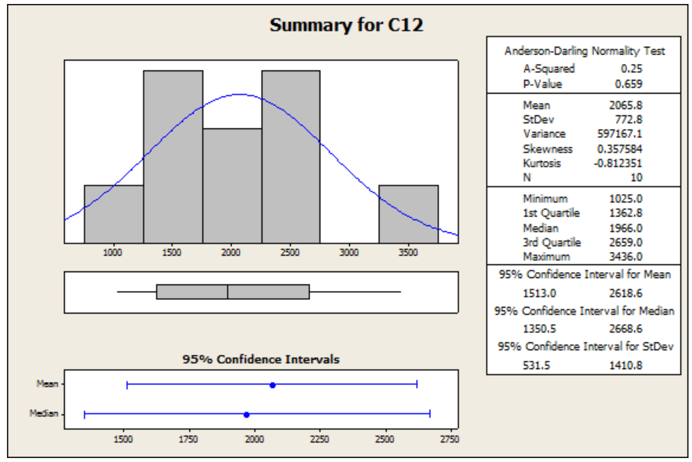


Fig. 16. Graphical Summary for Experiment C5 Which had the Best Results

-	True	False
Positive	14878	132622
Negative	127028	15476

TABLE XII
% CORRECT POS: 0.4901, % CORRECT NEG: 0.4892, OVERALL % ACCURACY: 0.4893

D. *Comparisons between Techniques*

The top performing experiment set from each training technique (based on mean analysis) were selected to be compared to one another to determine which training technique had the best results.

For all analysis, C1 will be raining technique one, C2 is training technique two, and C3 is training technique three.

1) Miami Image: The ANOVA comparison between these results can be seen in figure 17. The ANOVA analysis shows that experiment sets C1 and C2 do not have any difference between their results. One can also see that there is a substantial difference between C3 and the other experiments. A paired t-test was done 18 which shows that there is a definite statistical significant difference in their results.

2) Crinoid Image: An ANOVA comparison was done to observe the differences in the best performing experiment sets. Figure 19 shows these results. Just like the boat image analysis, there is no big difference between experiment sets C1 and C2 but there is a significant difference between these

One-way ANOVA: C1, C2, C3

```
Source  DF        SS        MS      F      P
Factor   2  65875552  32937776  29.39  0.000
Error   27  30260854   1120772
Total   29  96136406

S = 1059   R-Sq = 68.52%   R-Sq(adj) = 66.19%

                          Individual 95% CIs For Mean Based on Pooled StDev
Level   N  Mean  StDev  +---------+---------+---------+---------
C1     10  5729   1218                        (---*----)
C2     10  4676   1239                 (---*----)
C3     10  2195    587  (----*---)
                        +---------+---------+---------+---------
                     1500      3000      4500      6000

Pooled StDev = 1059
```

Fig. 17. ANOVA comparison between training techniques on Miami Image

Paired T-Test and CI: C2, C3

```
Paired T for C2 - C3

             N  Mean  StDev  SE Mean
C2          10  4676   1239      392
C3          10  2195    587      186
Difference  10  2482   1516      479


95% CI for mean difference: (1397, 3566)
T-Test of mean difference = 0 (vs not = 0): T-Value = 5.18  P-Value = 0.001
```

Fig. 18. paired t-test between top two training techniques on Miami Image

One-way ANOVA: C1, C2, C3

```
Source  DF        SS       MS     F      P
Factor   2   3978957  1989479  6.07  0.007
Error   27   8847421   327682
Total   29  12826379

S = 572.4   R-Sq = 31.02%   R-Sq(adj) = 25.91%


                            Individual 95% CIs For Mean Based on
                            Pooled StDev
Level   N    Mean  StDev  -------+---------+---------+---------+--
C1     10  2384.9  522.9                 (---------*--------)
C2     10  2512.2  563.6                     (--------*--------)
C3     10  1683.9  626.1  (--------*--------)
                          -------+---------+---------+---------+--
                              1600      2000      2400      2800

Pooled StDev = 572.4
```

Fig. 19. ANOVA comparison between training techniques on Crinoid Image

experiments and C3 (paired t-test seen in figure 20).

3) Results: These experiments show that the first two training techniques did not perform well. This suggests that a single 80x80 picture was not sufficient for a chromosome to learn off of. This meaning the larger training set would be needed. The results seen for the third training technique correlate with this idea because all experiments performed much better and there was a much larger training set.

E. Comparison with Previous Work

For a comparison between the results obtained in this work only the best performing experiment set (from the third training technique) for each picture will be compared to the best performing experiment set from the previous work (Assignment 2 [7]).

For both analysis, C1 corresponds to the best result obtained in this work and C2 is the best result from assignment 2.

1) Miami Image: The ANOVA results are seen in figure 21 and the paired t-test is in figure 22. these results show that the results obtained in this work is much better than the previous work but there is a problem. The language used is slightly different and the alpha value used in Poli's fitness update strategy was substantially different so any results showing that this work's results were better are not statistically valid as the experiments were not entirely fare. This is not a big concern because in the assignment 2 the Miami image was only used for learning system parameters to be used on the crinoid image and even then the results obtained in assignment 2 were already good for finding the boats. A comparison of the % accuracy of the best results from the two works can be compared. The best result in this work got an accuracy of roughly 95.94% and the best from assignment 2 was roughly 92.65%. This shows that this new training technique did produce better results.

Paired T-Test and CI: C2, C3

```
Paired T for C2 - C3

             N  Mean  StDev  SE Mean
C2          10  2512    564      178
C3          10  1684    626      198
Difference  10   828    666      211


95% CI for mean difference: (352, 1305)
T-Test of mean difference = 0 (vs not = 0): T-Value = 3.93  P-Value = 0.003
```

Fig. 20. paired t-test between top two training techniques on Crinoid Image

One-way ANOVA: C1, C2

```
Source  DF          SS          MS       F      P
Factor   1  8208987324  8208987324  538.90  0.000
Error   18   274193805    15232989
Total   19  8483181129

S = 3903   R-Sq = 96.77%   R-Sq(adj) = 96.59%


                           Individual 95% CIs For Mean Based on
                           Pooled StDev
Level   N   Mean  StDev  +---------+---------+---------+---------
C1     10   2195    587  (-*-)
C2     10  42714   5488                                 (--*-)
                         +---------+---------+---------+---------
                         0     12000     24000     36000

Pooled StDev = 3903
```

Fig. 21. ANOVA comparison between best results in both works on the Miami Image

Paired T-Test and CI: C2, C1

```
Paired T for C2 - C1

             N   Mean  StDev  SE Mean
C2          10  42714   5488     1736
C1          10   2195    587      186
Difference  10  40519   5627     1779


95% CI for mean difference: (36494, 44544)
T-Test of mean difference = 0 (vs not = 0): T-Value = 22.77  P-Value = 0.000
```

Fig. 22. paired t-test between best results in both works on the Miami Image

2) Crinoid Image: The ANOVA results are seen in figure 23 and the paired t-test is in figure 24. Once again, these results show that the results obtained in this work is much better than the previous work. There is again, a problem with this comparison. The language used was slightly different and there were two completely different fitness update strategies used in these two works. Assignment 2 used a basic weighted sum and this work used Poli's fitness update strategy. Comparing these fitness values is not fare as they do not directly correspond to how well the GP system classified the results (as mentioned earlier). We can compare the % accuracy though, and the best result in this work got a best of roughly 48.93% and the best from assignment 2 was roughly 73.269%. This shows that there is not an improvement with this training technique on the crinoid image.

F. Equivalent System Parameter Results and Comparison with Previous Work

Although there was an attempt to keep the system parameters the same a lot of variables were introduced (such as the fitness function). Another set of experiments were done that had all system parameters identical to the large set of comparisons done in assignment 2. This set of experiments had 30 runs done to ensure statistical significance.

One-way ANOVA: C1, C2

```
Source  DF          SS          MS        F      P
Factor   1  1463794720  1463794720  3402.87  0.000
Error   18     7742970      430165
Total   19  1471537690

S = 655.9   R-Sq = 99.47%   R-Sq(adj) = 99.44%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N   Mean  StDev  -------+---------+---------+---------+--
C1     10   1684    626  (*)
C2     10  18794    684                                     (*
                         -------+---------+---------+---------+--
                              5000     10000     15000     20000

Pooled StDev = 656
```

Fig. 23. ANOVA comparison between best results in both works on the Crinoid Image

Paired T-Test and CI: C1, C2

```
Paired T for C1 - C2

             N    Mean  StDev  SE Mean
C1          10    1684    626      198
C2          10   18794    684      216
Difference  10  -17110   1147      363

95% CI for mean difference: (-17931, -16290)
T-Test of mean difference = 0 (vs not = 0): T-Value = -47.18  P-Value = 0.000
```

Fig. 24. paired t-test between best results in both works on the Crinoid Image

This set of experiments had an identical language, same fitness function and same number of generations and population size. The other system parameters used for this set were based off of and were set to be close to the system parameters for the best performing set of experiments in assignment 2. The only thing that is different between these two sets of experiments is the version number of the GP system used. This set of experiments were done on the newest version of “linkable GP”. It is not expected that this will affect the results in any way.

There were four experiments done in this suit. Their system parameters can be seen in table XIII. These system parameters are unique but are similar to the system parameters of the best performing experiment set from assignment 2.

The ANOVA results for this group of experiments can be seen in figure 25.All of these results are similar to one another but it is clear that the best performing one (based on mean value) is experiment C3. A paired t-test was done between C3 and C1 (the next best performing experimental run) which showed that there is no actual statistical difference between these results as the p-value obtained was 0.54.

Figure 26 is the graphical summary of experiment set C5.

-	Gens	Pop	Cross	Mut	Min	Max	Radius	Weight
C1	20	50	.5	.7	5	15	2	5.5
C2	20	50	.7	.7	5	15	5	5.5
C3	20	50	.8	.8	5	15	3	5.5
C4	20	50	.65	.65	5	15	3	5.5

TABLE XIII
SYSTEM PARAMETERS FOR CRINOID IMAGE ANALYSIS WITH TRAINING TECHNIQUE TWO (NOTE: MIN AND MAX REFER TO TREE DEPTH AND WEIGHT IS THE BONUS THE CHROMOSOME GETS FOR CLASSIFYING A PIXEL IN A FOSSIL CORRECTLY)

One-way ANOVA: C1, C2, C3, C4

```
Source   DF         SS       MS     F      P
Factor    3    8266856  2755619  2.86  0.039
Error   152  146275783   962341
Total   155  154542639

S = 981.0   R-Sq = 5.35%   R-Sq(adj) = 3.48%

                             Individual 95% CIs For Mean Based on
                             Pooled StDev
Level   N    Mean   StDev  ---+---------+---------+---------+------
C1     39  3048.0   889.0      (----------*---------)
C2     39  3482.3  1054.0                     (---------*---------)
C3     39  2908.9  1055.7  (---------*---------)
C4     39  3356.3   913.1                (---------*---------)
                           ---+---------+---------+---------+------
                           2700      3000      3300      3600

Pooled StDev = 981.0
```

Fig. 25. ANOVA results with training technique 3 and 30 experimental runs on the Crinoid image.

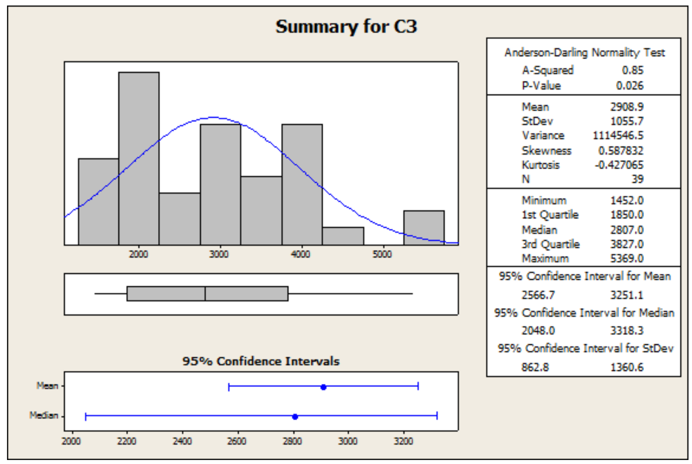


Fig. 26. Graphical Summary for Experiment C3

There is no normal distribution in this graph experiment set and the standard deviation and 90% confidence interval is also pretty high. Unlike the other experiments that used Poli’s fitness function (All learning curves showed the same trend and average of average could not be plotted due to the nature of the fitness function), this one used a basic weighted sum so a learning curve could be included that makes sense. Figure 27 shows the learning curve.

The language obtained in the best performing experiment set in C6 is as follows(simplified):

```
(avgRed) > (avgGrey/(12.02/stdDev))
```

Note that the divide is a protected divide (does not allow divide by zero errors).

The confusion matrix can be seen in table XIV.

A proper comparison can now be done between the two works. The ANOVA comparison (figure 28) between the best

-	True	False
Positive	13307	80339
Negative	179311	17047

TABLE XIV
% CORRECT POS: 0.4383, % CORRECT NEG: 0.6905, OVERALL % ACCURACY: 0.6641

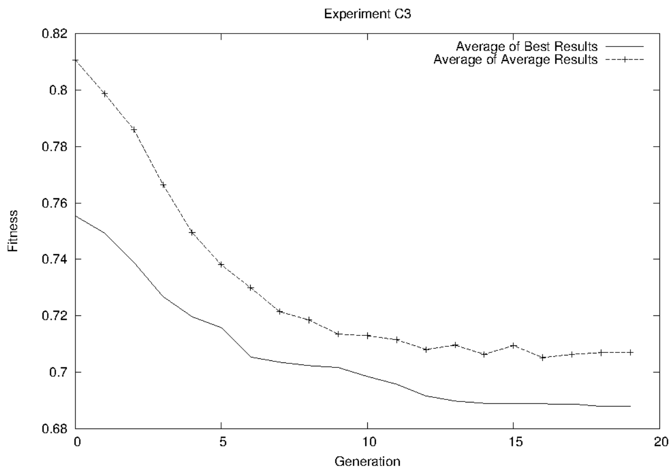


Fig. 27. Learning curve for Experiment C3

```
One-way ANOVA: C1, C2

Source  DF          SS          MS        F      P
Factor   1  3591701088  3591701088  3639.82  0.000
Error   58    57233189      986779
Total   59  3648934277

S = 993.4   R-Sq = 98.43%   R-Sq(adj) = 98.40%

                               Individual 95% CIs For Mean Based on
                               Pooled StDev
Level   N   Mean  StDev  -----+---------+---------+---------+----
C1     30   3086   1064  (*)
C2     30  18560    917                                (*)
                         -----+---------+---------+---------+----
                           5000     10000     15000     20000

Pooled StDev = 993
```

Fig. 28. ANOVA results comparing the best results form the two works on the Crinoid image.

performing experiment in this set (C3) is now considered C1, and the best preforming experiment from assignment 2 is called C2.

The ANOVAs p-value indicates that there is a considerable statistical difference between these two experiments (30 runs were done on these two so statistical significance could be shown). Because the only difference between these two experiments was the training technique it suggests that the training technique used in this work was more effective (training technique three was used).

VI. Conclusions and Future Work

The results obtained in this work suggest that both training technique one and training technique two are not sufficient for GP system training for these images. Training technique three obtained the best results from all of the experiments done in this work. Training techniques one and two both only allowed the GP to train on a single 80x80 pixel image for each chromosome per generation where training technique three allowed for all five 80x80 training images to be given to a chromosome per generation. This suggests that the GP system needs more of the whole image to train on in order to obtain results.

When comparing the results obtained from the early experiments in this work to the previous work one can see that the fitness values obtained from this work were substantially better but unfortunately the fitness value does not directly correspond to a better classifying language. This is because the fitness value is not directly related to the number of correct classifications; instead there the number of correct classifications is put through a mathematical equation 1. When studying the % accuracy we can see that this work got better results for the Miami image, but not for the crinoid image.

On the last major set of experimental analysis that was done between the results obtained in this work and assignment 2 it was shown that training technique 3 from this work allowed the system to obtain much better fitness values. These comparisons were done in a way that allowed little to no variables into the system. That being said, the training used in assignment allowed the system to train on 24459 pixels and training technique 3 in this work allowed it to train on 32000 (7541 more pixels)(This was only true for half of the generations, once ten generations completed, training technique 3 only trained on 19200 pixels). The fact that the chromosomes were able to see more pixels in the beginning half of the training (in this work) could be the only reason that it performed better.

There are still some problem areas in the classification images. These problem areas are the same as in assignment 2. These problem areas are white areas on the land for the boat images and other fossilized organisms with similar textures to crinoids in the crinoid image. For both images the edges of the objects being classified were problem areas but problems could probably be ignored as one would simply look at a classified image and know that there is an object in that area.

Future works that could be derived from this work would be to train on other images and try to classify other objects with GP image analysis. It would also be interesting to try training technique three again with larger images than 80x80 pixels.

References

[1] Bailey, A. *"Evolving a Spacial Image Analyzer for Tree Detection Using Fitness Sharing"* St. Catharines, ON: Brock University, Computer Science Department.

[2] Bengtsson, P. *"Genetic Programming Applied to an Image Analysis Problem"* Uppsala, Sweden: Uppsala University, Computer Science Program.

[3] Espejo, P.G. Ventura, S. Herrera, F. "A survey on the application of genetic programming to classification" *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 40.2 (2010): 121-144.

[4] F. Fernandez-de Vega, G. Gil, J. Gomez Pulido, and J. Guisado. "Control of bloat in genetic programming by means of the island model". In *Parallel Problem Solving from Nature-PPSN VIII*, pages 263271. Springer, 2004.

[5] Holland,J. "Adaptation in natural and artificial systems", University of Michigan press Ann Arbor, MI, 1975, 1, 5.

[6] Hughes, J.A. (2012) *COSC 4V82 Assignment 1 - "Exploring Genetic Programming and its Ability to Learn Different Problem Sets"* St. Catharines, ON: Brock University, Computer Science Department.

[7] Hughes, J.A. (2012) *COSC 4V82 Assignment 2 -"Applying Genetic Programming to Image Analysis"* St. Catharines, ON: Brock University, Computer Science Department.

[8] Medland, M. (2012) *Linkable GP* St. Cathatines, ON: Brock University, Computer Science Department Internet: https://bitbucket.org/mmedland/linkablegp/overview [November 16, 2012].

[9] Poli, R. *"Genetic Programming for Image Analysis"* Birmingham, UK: The University of Birmingham, School of Computer Science.

[10] Dr. Ross, B. “Brock COSC 4V82 Genetic Programming.” Internet: http://www.cosc.brocku.ca/Offerings/4V82/ [October 16, 2012].
[11] University of Pittsburgh ”Sandstones.” Internet: http://www.pitt.edu/ cejones/GeoImages/5SedimentaryRocks/Carbonates/Bioclastic.html [October 21, 2012].
[12] Viney, M. ”The Virtual Petrified Wood Museum.” Internet: http://petrifiedwoodmuseum.org/ [March 25, 2012].