A Comparison of Finite State Classifier and Mahalanobis-Taguchi System for Multivariate Pattern Recognition in Skin Cancer Detection

Elizabeth A. Cudney and Steven M. Corns

Abstract - This project presents two methods for image classification for the detection of malignant melanoma: the Mahalanobis-Taguchi System and Finite State Classifiers. The Mahalanobis-Taguchi System is a diagnosis and predictive method for analyzing patterns in multivariate cases, while Finite State Classifiers are a state based machine learning technique. The goal of this study is to compare the ability of the Mahalanobis-Taguchi System and a Finite State Classifier to discriminate using small data sets. We examine the discriminant ability as a function of data set size using publicly available skin lesion image data. While analysis of the data shows a high degree of correlation, the Mahalanobis-Taguchi System performed poorly when trying to discriminate between Malignant Melanoma and benign lesions. Alternately, the Finite State Classifiers developed using evolutionary computation obtained over 85% correct classification of the malignant and benign lesions using the image data sets.

I. Introduction

Malignant melanoma is the deadliest form of skin cancer. However, if treated during early stages, it has a good prognosis. Color analysis has shown to provide critical discriminating information in the diagnosis of melanoma [1]. Detection methods rely on guidelines such as asymmetry, border irregularity, color variegation, and diameter of the skin lesion. This research uses data collected from pigmented lesion images to investigate features that detect malignant melanoma. The study compares the discrimination ability of the Mahalanobis-Taguchi System to Finite State Classifiers.

A. Mahalanobis-Taguchi System

Mahalanobis distance (MD) was developed by well-known statistician P.C. Mahalanobis in 1932 to help identify members of a group defined by characteristics which may or may not be correlated. Dr. Genichi Taguchi led the development of the Mahalanobis-Taguchi System (MTS) by providing a means to define the reference group and measure the degree of abnormality of individual observations [2].

Mahalanobis distance is a distance measure based on correlations between variables and the different patterns that can be identified and analyzed with respect to a reference population. MD is a discriminant analysis approach, which is used to determine the similarity of a known set of values (normal group) to that of an unknown set of values (abnormal group). MD has successfully been applied to a broad range of cases mainly because it is very sensitive to inter-variable changes in data. Also, because the MD is measured in terms of standard deviations from the mean of the samples, it provides a statistical measure of how well an unknown sample matches a known sample set.

The Mahalanobis-Taguchi System is a pattern information technology, which has been used in different diagnostic applications to make quantitative decisions by constructing a multivariate measurement scale using data analytic methods. In MTS approach, MD (a multivariate measure) is used to measure the degree of abnormality of patterns and principles of Taguchi methods are used to evaluate accuracy of predictions based on the scale constructed. The advantage of MD is that it takes into consideration the correlations between the variables and this is very important in pattern analysis. MTS is a very economic approach for multidimensional pattern recognition systems.

A pattern is a definable lack of randomness. For example, a pattern could be a fingerprint image, a handwritten cursive word, a human face, or a speech signal. Pattern recognition is the study of how to observe and distinguish patterns of interest, and make reasonable classification decisions [3].

In multidimensional systems, it is an economic necessity to reduce the number of variables by neglecting those that have little or no effect on the measurement function. Numerous approaches have been conducted previously, such as linear discriminant analysis, logistic regression decision trees, and neural networks [4].

B. Finite State Classifiers

A Finite State Classifiers (FSC) is a finite state machine that has been modified for classification of data. Within each state an evaluation of the data is performed to drive the classifier. Based on this evaluation a response and a transition are generated, with the response a vote for how the data should be classified and the transition dictating the next state the classifier should use for its next evaluation. By summing all of the responses the overall classification of the data is determined.

This type of classifier is constructed using machine learning techniques. Training data is provided to the classifier and the parameters within the states are adjusted so that the finite state machine correctly classifies this training data. The ability of the classifiers found is then tested using a validation

Elizabeth A. Cudney is an Assistant Professor in the Engineering Management & Systems Engineering Department at Missouri University of Science & Technology, Rolla. MO, USA (e-mail: cudney@mst.edu).

Steven M. Corns is an Assistant Professor in the Engineering Management & Systems Engineering Department at Missouri University of Science & Technology, Rolla. MO, USA (corresponding author: e-mail: cornss@mst.edu).

978-1-4244-9897-0/11/$26.00 ©2011 IEEE

set of data. Ideally the finite state classifier will identify relationships within the data set that indicate group membership and store this information for future analysis.

The goal of this study is to compare the ability of MTS and finite state classifiers to detect skin cancer. Section 2 provides a review of relevant literature in MTS and Finite State Classifiers. Section 3 describes the details of MTS. Section 4 presents the Finite State Classifier algorithm. Section 5 shows a comparison study of MTS and FSC using pigmented lesion images to detect malignant melanoma. Section 6 summarizes the results and provides a conclusion on the research.

II. Background

Considerable research is available utilizing MD to determine similarities of values from known and unknown samples. Existing research uses the MTS for prediction and diagnosis illustrating the methodology"s accuracy and effectiveness. However, little is presented to compare the accuracy and effectiveness of the MTS to other methodologies.

Taguchi [5] developed the Mahalanobis-Taguchi System for diagnosis and pattern recognition using a case study involving liver disease diagnosis in Tokyo, Japan using fifteen variables. Dr. Taguchi developed an eight-step procedure entitled "Mahalanobis Distance for Diagnosis and Pattern Recognition System Optimization Procedure".

Mahalanobis distance is utilized to determine the misclassification of samples in the research by Shen et. al [6]. Two processes for tablet production are evaluated in this research including wet granulation and direct compression which are the two main methods used in tablet preparation. Pyrolysis-gas chromatography-mass spectrometry was used to discriminate the two processes. First, samples were removed that did not contain at least half the number of samples for one of the two classes. Principal components analysis (PCA) was then employed to determine the main principal components. Based on the PCA analysis, three factors and one sample were excluded. The data was further processed using Fischer discriminant analysis to classify the sample. Fisher discriminant analysis is an approach is a commonly used for feature extraction between groups [7]. The results were evaluated using MD to determine the misclassification rate.

Mahalanobis distance was also used to maximize productivity in a new manufacturing control system by Hayashi et al. [8]. The research used MD as a core to their manufacturing control system because of the method"s ability to recognize patterns. The new system detected deviations from normal productivity much earlier and enabled root cause identification and prioritized resolution.

Al-Otum [9] proposed two algorithms for color morphology. The first is a Mahalanobis-color-distance-based morphological ordering algorithm. The second is a corrected component wise morphological ordering algorithm. Both algorithms employ a Mahalanobis color measure based on reduced and conditional ordering of the data to perform four basic morphological operators: dilation, erosion, opening, and closing. The data ordering is performed using the Mahalanobis-color-distance (MCD) in which the angle-valued pixels are replaced by a scalar. The two algorithms were evaluated using a perceptual image quality assessment.

Asada [10] used MTS to forecast the yield of wafers. Yield of wafers is determined by the variability of electrical characteristics and dust. The research focused on one wafer product that had a high yield. MDs were calculated on various silicon wafers to compare the relationship between yield and distance. The signal-to-noise ratios were used to indicate the capability of forecasting and the effect of the parameters. This research showed the applicability of the MD to predict the defective components.

Garcia-Lagos et al. [11] utilized Mahalanobis distance in a topology assessment for power systems. The research developed system architecture for use in a state estimator which worked with a bus-branch oriented network model. The architecture was developed in two stages including a pre-processing stage and a classification stage. The pre-processing stage transformed the measurements into output vectors for grid topologies. In the second stage, classification is determined using a layer of Gaussian potential function units based on the MD. The units are established through input from the pre-processed vector. This indicates the degree to which a vector belongs to a specific topology. This assessment enabled researchers to identify the actual topology.

Pattern recognitions using MD was demonstrated in the work of Wu [12]. In this research, pattern recognition was used to diagnose human health. The results of tests from a regular physical check-up were used as the characteristics. The correlation between different tests was shown. MS was used to summarize the multi-dimensional characteristics into one scale. In this research the base point was difficult to define since it was a healthy person. People who were judged to be healthy for the past two years were considered to be healthy. The research considered diagnosis of liver function with the objective to forecast serious disease until the next check-up. The approach provided a more efficient method that also avoided inhuman treatment that had previously used double blind tests.

Jugulum and Monplaisir [13] performed preliminary comparison between MTS and neural networks. They used medical data with 15 variables. They compared both methods for the large sample and small sample. The small sample was selected from the complete sample of 200 observations in the healthy group. They showed that in the case of large samples both methods performed equally well and in the case of small samples, MTS is somewhat better than neural networks. In this research, they did not compare these methods in terms of reducing the number of attributes.

The research presented here is distinct from existing literature. The research conducted in this paper compares the accuracy and effectiveness of the MTS and Finite state classifiers to neural networks for different data sizes for the first time.

Finite state classifiers are a specific application of finite state machines to solve a classification problem. Finite state machines are a standard representation for behavior in

artificial intelligence problems dating back to work by Larry Fogel [14]. They have been used in studying the prisoner‟s dilemma problem for several years, including work by David Fogel [15] and Stanley, Ashlock, and Tesfatsion [16]. Chellapilla and Czarnecki [17] expanded the capabilities of finite state machines by introducing the concept of modularity, greatly increasing the rate at which optimal control concepts were found.

The use of finite state machines for classification has been done primarily on the classification of Polymerase Chain Reaction (PCR) primers for the maize by Ashlock, et al. [18]. There have been several refinements to this method. One such modification was the use of hybridization to blend high performance state machines to develop classifiers with even higher performance [19]. This work was expanded to analyze PCR primers for mice by Yadav and Corns [20]. Based on the success of these applications, this work will attempt to expand that work to include classification of image data for medical diagnosis.

III. Experimental Design

In the Mahalanobis-Taguchi System, the Mahalanobis space (MS) is obtained using the standardized variables of healthy or normal data. The MS can be used to discriminate between normal and abnormal objects. Once this MS is established, the number of attributes is reduced using an orthogonal array (OA) and signal-to-noise ratio (SN) by evaluating the contribution of each attribute. Each row of the OA determines a subset of the original system by including and excluding each attribute of the system. The different stages of MTS method are summarized below:

Stage I: Construction of a Measurement Scale

- Select a reference group with suitable variables and observations that are as uniform as possible.
- Use this reference group as a base or reference point of the scale.

Stage II: Validation of the Measurement Scale

- Identify the conditions outside the reference group.
- Compute the Mahalanobis distances of these conditions and check if they match with decision-maker‟s judgment.
- Calculate S/N ratios to determine the accuracy of the scale.

Stage III: Identify the Useful Variables (Developing Stage)

- Determine the useful set of variables using orthogonal arrays and signal-to-noise ratios.

Stage IV: Future Diagnosis with Useful Variables

Monitor the conditions using the scale, which is developed with the help of the useful set of variables. Based on the values of the Mahalanobis Distances, appropriate corrective actions can be taken.

The general procedure of MTS is described as follows [2]. The first step in MTS is to construct a measurement scale using the MS as a reference. To construct a measurement scale, a data set of the normal observations needs to be collected. The collected normal observations are then standardized using Equation 1.

$$Z_i = \frac{X_i - m}{\sigma} \tag{1}$$

where,

- m, mean of the attribute,
- σ, standard deviation of the attribute,
- Z_i, standardized variables, and
- X_i, normal observations.

The standardized vector is obtained from the standardized values of X_i (i = 1, 2, …, k). MD measures the distance in multidimensional spaces by accounting for the correlation among the attributes. The statistical meaning of MD is the nearness of an unknown point to the mean of the group. Equation 2 is the formula used to calculate MDs.

$$MD_j = D_j^2 = \frac{1}{k} Z_{ij}^T C^{-1} Z_{ij} \tag{2}$$

Where C^{-1} is the inverse of the correlation matrix which contains correlation coefficients between the variables, T is the transpose of the standard vector, and k is the number of data sets. It can be easily proved that the average value of the MDs is 1 for all the observations in the MS. For this reason, MS is also called the unit space [2].

The second step is to validate the measurement scale. In order to validate the measurement scale, observations outside of MS are used, usually abnormal or test observations. The mean value, standard deviation and correlation matrix of the normal observations are used to calculate the MD of the abnormal observations. For good measurement scales, the MDs of the abnormal observations are larger than the MDs of the normal observations.

The third step of MTS is to optimize the system. For this purpose, orthogonal arrays (OA) and signal-to-noise array (SN) are very useful to identify which attributes are important. In the experiment, every factor is assigned to a column in the OA, and every row represents the experimental combination of a run. A two level OA is used to represent inclusive (presence of a variable) and exclusive (absence of a variable). In a two level OA, 1 and 2 indicate the level corresponding to the presence and absence of the variable, respectively. Each attribute will be used or neglected with respect to the OA and the SN ratio is calculated.

Various types of SN ratio exist; however, MTS uses the larger-the-better (LTB) or dynamic SN ratio. In the context of MTS, SN ratio is defined as the measure of accuracy of prediction of the scale. It reflects the severity of the abnormalities and the difference of the average SN values of each attribute when the attribute included and excluded. The classification ability is compared with the feedforward artificial neural network. In the aspect of data size, efficiency and time, MTS performs well when compared to neural

network. Equations 3 and 4 give the dynamic and the larger-the-better (LTB) SN ratio, respectively.

$$DynamicSN = \eta = 10\log\left(\frac{\frac{1}{r}\left(S_\beta - V_e\right)}{V_e}\right) \quad (3)$$

where,

- S_T = total sum of squares

$$S_T = \sum_{i=1}^{t} y_i^2$$

- r = sum of squares due to input signal

$$r = \sum_{i=1}^{t} M_i^2$$

- S_β = sum of squares due to slope

$$S_\beta = \frac{1}{r}\sum_{i=1}^{t}\left(M_i y_i\right)^2$$

- S_e = error sum of squares

$$S_e = S_T - S_\beta$$

- V_e = error variance

$$V_e = \frac{S_e}{t-1}$$

$$LTBSN = \eta = -10\log_{10}\left[\frac{1}{t}\sum_{i=1}^{t}\left(\frac{1}{D_i^2}\right)\right] \quad (4)$$

Where D_1^2, D_2^2, …, D_t^2 are distances corresponding to the abnormal situations.

For a given attribute X_i, SN^+ represents the average SN ratio of including the attribute X_i. SN^- represents the average SN ratio when the attribute X_i is excluded. The gain is the difference between SN^+ and SN^- as shown in Equation 5.

$$Gain = SN^+ - SN^- \quad (5)$$

If the gain is positive, the recommendation is to use the attribute (i.e., the attribute is considered "useful"), if not the attribute is neglected. A confirmation run is performed by constructing an MS with the useful variables. The MDs of the abnormal observations are also calculated based on the set of useful variables. The average MD of the normal group is compared to the average MD of the abnormal group.

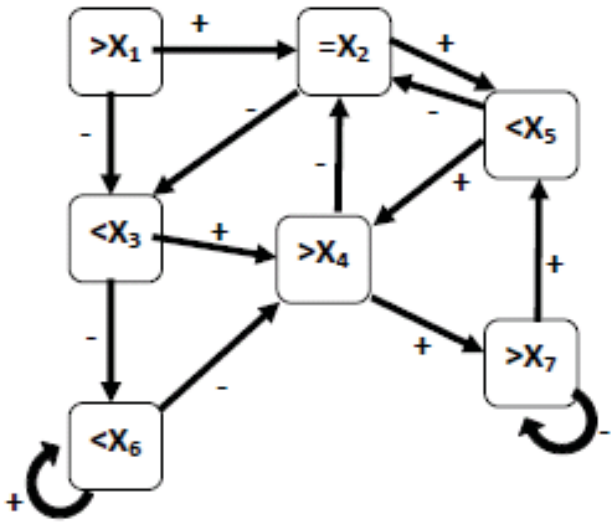

Figure 1, Finite State Classifier similar to that used in this study. X values represent evolved values being compared to image data, and the arrows represent transitions and their accompanying response

The second portion of this study is the implementation of a machine learning system to create the finite state classifiers. For this work an evolutionary computation approach was used to create these classifiers. Two key features of any evolutionary computation based machine learning system are the data structure holding the problem solutions to be evolved and the fitness function. The data structure used in this project is a finite state machine with varying states and transitions driven by the results of image data. Figure 1 shows part of such a machine. The states of the FSC have three possible types: ? (don't know), + (positive for melanoma), and - (negative for melanoma). These state labels are used to permit the finite state machine to function as a classifier, though not in the usual manner, as explained when the fitness function is specified. There is some evidence that that this type of evolutionary methods can handle some problem complexity better than these other representations [15], although that has not been shown when using real valued information with this type of FSC.

For training a population of FSCs we used a collection image samples. The fitness of a FSC on a training set of image data is computed as follows. Each image data relating to the total color values (mean and max values for RGB) and the shape of the lesion is run through the FSC to allow all of the image data to be analyzed (17 states). As the machine passes through each state, values from a payout table are summed. These numbers represent complete neutrality to all factors except "positive for melanoma" and "negative for melanoma." Table 1 shows the labels for the relationship between whether the sample is positive or negative for melanoma and the available states. The two values labeled with a W are incorrect classification of the sample, C designates correct classification, and the D indicates that the state is a „don't know". Fitness for an FSC is summed over all image data sets in the training set.

TABLE 1, PAYOUT MATRIX LABELS

	+	-	?
Melanoma	C	W	D
Non-Melanoma	W	C	D

This fitness function rewards the FSC incrementally, after each state transition. This incremental fitness permits the FSCs to be indecisive, allowing for flexibility in sample evaluation. Large positive scores indicate a sample that the FSC classifies as quite likely to be positive for melanoma; large negative scores are votes that the sample is of a non-cancerous lesion. Scores near zero, however, indicate either ignorance (the image does not match existing information) or confusion (the image matches features of both positive and negative melanoma images).

IV. RESULTS AND DISCUSSION

The purpose of the research was to compare the discriminating ability of MTS to Finite State Classifiers, and then compare these results to neural network results. The comparison study utilized pigmented lesion images to detect malignant melanoma. The data consisted of seventeen independent variables and one dependent variable. The data consisted of 1,832 data sets. Of these data sets, 1,683 were considered normal or benign and 149 were considered malignant.

A. MTS Analysis

The study uses data collected from pigmented lesion images to investigate features that detect malignant melanoma [21]. The goal is to predict whether a sample is malignant or benign. Separate benign and malignant data sets are then used to calculate the MS. The benign data consisted of 1,683 samples and 149 data sets are malignant. The healthy (benign) data sets are generalized. The correlation matrix and inverse correlation matrix are calculated. Finally, the Mahalanobis distances for the selected data sets are calculated. The correlation matrix, standard deviation and mean of the healthy data sets are used for the MD of the entire data set. In the case of the normal (healthy) data sets, the MD value is very small and the average of MD is close to one. The abnormal MD values are larger than normal which illustrates the classification ability of MD.

The next step is to optimize the system. An L32 orthogonal array was used for the optimization. The useful set of variables is obtained by evaluating the gain, measured in decibels, associated with each factor. Gain is the average difference in the magnitudes of signals using the S/N ratio measured in decibels when a variable is excluded and included. The gains for the 17 factors are shown in Fig. 2, 13 of which are useful.

Regression analysis is then performed. The useful variables are used in the regression analysis. Using the regression equation for the 13 factors determined to be useful variables, the predicted percent overlap as a measure of benign

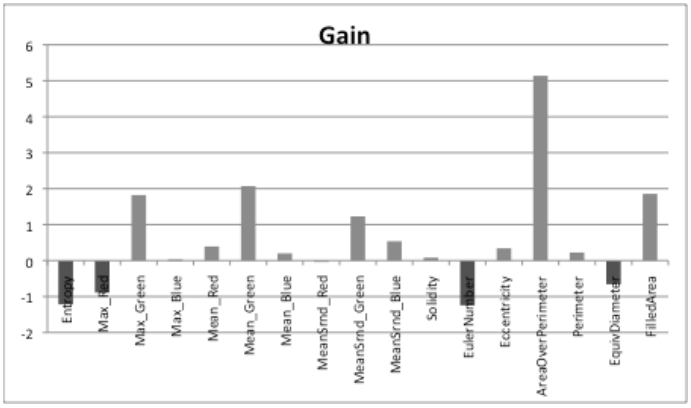

Figure 2, Gains for the MTS method based on skin image data.

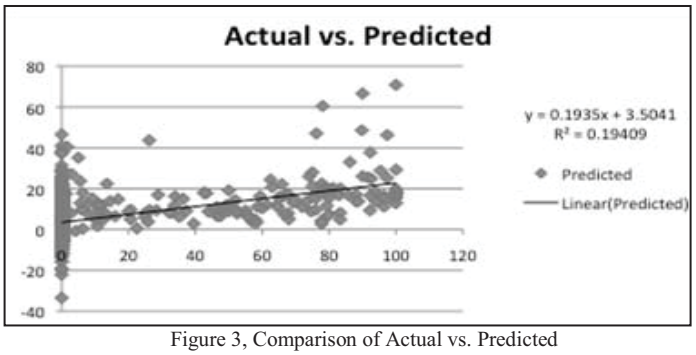

Figure 3, Comparison of Actual vs. Predicted

compared to the actual values graphically. Scatter plots were then developed to show the correlation between the measured versus predicted values (Fig. 3). The MTS method was employed for the comparison study because of its reported benefits with small, highly correlated data sets. However, the prediction capability for this data set provided a correlation between the actual and predicted values of 0.44.

B. FSC Analysis

The FSC experiments used the same data as the MTS analysis, although it was used differently to match the requirements of the machine learning method used. Because the data set contained only 149 samples that were positive for malignant melanoma, 149 of the other samples were randomly selected and mixed with the positive samples to provide a balanced data set in hopes of reducing bias. This data was then separated into sets of image data for training and 98 sets of image data for testing.

Six experiments were performed, three using a population of 30 FSCs and three with a population size of 600 FSCs. Three of the experiments used FSCs with 64 states, and the three remaining used 128 states. The FSCs are all initialized by filling in both transitions and state labels at random. In addition, each state has a variable ranging from zero to one and a variable to indicate an operator for comparing values {<, = , >}. For all experiments single tournament selection with tournament size four was used. The population is randomly shuffled into groups of four, with the two most fit FSCs

TABLE 2, FSC EXPERIMENTS EXAMINED (100 REPLICATES EACH)

	Population Size	Generations	Number of States
Experiment 1	600	1000	64
Experiment 2	600	1000	128
Experiment 3	600	5000	128
Experiment 4	30	1000	64
Experiment 5	30	1000	128
Experiment 6	30	5000	64

TABLE 3, BASELINE PAYOUT MATRIX

	+	-	?
Good	1	-1	0
Bad	-1	1	0

reproducing and replacing the two least fit in each group with their children. Reproduction treats the string of states in an FSC as a linear chromosome. The two FSCs reproducing are copied, the copies undergo two-point crossover, and then each copy is subjected to one mutation. The mutation modifies the initial state of the FSC 10% of the time, randomly changes the comparison operator 10% of the time, randomly picks a new destination for one of the transitions 20% of the time, and modifies the label {+,-, ?} on a state 60% of the time. During crossover, the initial state of the FSC moves with the first state. One hundred simulations with distinct starting populations were performed, saving the best FSC from each simulation.

The six experiments used to analyze the finite state classifier methods are given in table 2. For these experiments 1000 or 5000 generations were used to evolve the FSCs and the payout table shown in table 3 was used to assigned fitness values, which were then evaluated using the cross validation data. The population size, number of generations, and number of states were varied to explore parameter setting in hopes of increasing performance. For each of the experiments, 100 runs were conducted and the classifier that did best on the cross validation data from each experimental run was saved.

The FSC experiments met with varied success, ranging from below 50% correct classification rate to almost 88% correct classification rate. The first experiment had the best individual FSC, with the best classifier only failing to correctly classify 12 image data sets and one undecided (87.6% correct, table 4). Four other FSCs had classification success rates of greater than 84%. Unfortunately, most of the errors in classification were improper classification of malignant melanoma as benign lesions. In addition, four of the FSCs found performed worse than classifying all data sets as positive (more than 46 errors).

Of the other experiments, only the fourth experiment had results comparable to experiment 1, with the best FSC found providing one false positive more than the best FSC from experiment 1. It should also be noted that the worst performing classifiers from experiment 4 were an improvement over the worst of experiment 1, with only one

TABLE 4, PREDICTIONS VERSUS TRUTH RESULTS FOR THE MOST FIT FSC LOCATED IN EXPERIMENT 1

	+	-	?
Positive	42	9	1
Negative	3	43	0

TABLE 5, PREDICTIONS VERSUS TRUTH RESULTS FOR THE MOST FIT FSC LOCATED IN EXPERIMENT 4

	+	-	?
Positive	42	9	1
Negative	4	42	0

mis-classifying more than forty data sets and based on this small sample size a lower average number of errors (26.87) than experiment one (27.54 errors). When the number of generations was increased fivefold for experiment six, there was no improvement in the best FSC, and an increase in the average number of errors (to 27.27 errors), most likely due to overtraining on the first data set.

When the number of states was increased to 128 for experiments two, three, and five, the performance of the FSCs broke down. About 70% of all FSCs were a variation of „always positive“ or „always negative“. Many of the others were similar in performance, with most being no better than using a fair coin. The number of FSCs that performed better than this were three, six, and five for experiments two, three, and five respectively.

V. CONCLUSION AND FUTURE WORK

This research investigated how data correlations in the image data sets for skin lesions might be used as an indicator of malignant melanoma. The image information can be readily analyzed and shown to be highly correlated, but the MTS technique performed poorly when applied to this problem. One possible cause of this could be the large number of different colorations that occur on human skin which could be introducing additional variance and mixing different correlations within the data set, preventing the MTS approach from capturing the proper associations. This can be seen in the number of abnormal data points found during the MTS analysis. Further analysis of the data set could indicate possible interferences and improve this methods performance.

The Finite State Classifiers performed well on this problem, with the best classifiers correctly identifying almost 88% of the lesions. This is a notable improvement when compared to previous techniques, such as neural networks [21] and gradient vector flow snakes [22] which achieved success rates of no more than 86% and 80% respectively. The results for the experiments using sixty-four states were encouraging, although repeating the experiments using a larger sample size would validate the use of this method. Increasing the number of generations had a minor negative effect on the progress of the FSCs, Regardless of the population size and number of generations, when the number of states was increased to one hundred and twenty-eight the progress of the FSCs stalled and most of the classifiers performed no better than a fair coin.

This is likely due to the larger number of states delaying progress, but the impact was larger than expected.

The ability to quickly and easily diagnose malignant melanoma is important as the incident rate for this cancer has more than tripled for light skinned races in the past 20 years and is the sixth most common form of cancer in the United States [23]. Early detection and treatment gives the highest success rate in curing the disease, making methods capable of screening based on a simple picture of the lesion beneficial in that it is easy, painless, and fast for the patient. As with any method that does not give an extremely high accuracy a FSC method should only be used as a pre-screen and should not be used to replace other screening methods. Another interesting area of study would be to introduce a bias towards predicting malignant tumors. While the number of errors may increase, it would increase the usability of the tool so that it errors on the side of caution.

While further analysis is necessary, the development of a hybrid method between the MTS analysis and a machine learning technique like FSCs could potentially be of benefit.. It is thought that MTS was confounded by contradicting or superimposed image data (ie. a freckle on a mole). If this information could be removed using machine learning a stronger screening tool could be developed.

REFERENCES

[1] Faziloglu, Y., Stanley, R., Moss, R., Van Stoecker, W., and McLean, R., "Colour Histogram Analysis for Melanoma Discrimination in Clinical Images," *Skin Research and Technology*, 9, 147-155, 2003.

[2] Taguchi, G and R. Jugulum, "New Trends in Multivariate Diagnosis", *Indian Journal of Statistics*, 62, Series B, 2 233-248 (2000).

[3] Taguchi, G and R. Jugulum, The Mahalanobis-Taguchi Strategy: A Pattern Technology System, John Wiley & Sons, Inc., 2002.

[4] Jain, A.K., Duin, R.P.W., and Mao, J., "Statistical Pattern Recognition: A Review", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, Jan. 2000.

[5] Taguchi, S., "Mahalanobis Taguchi System", *ASI Taguchi Symposium*, 2000.

[6] Shen, H., Carter, J.F., Brereton, R.G., and C. Eckers, "Discrimination Between Tablet Production Methods Using Pyrolysis-Gas Chromatography-Mass Spectrometry and Pattern Recognition," *Analyst*, 128 (3), 287-292, 2003.

[7] Christensen, R., Analysis of Variance, Design and Regression: Applied Statistical Methods, Chapman & Hall/CRC, New York, NY, 1998.

[8] Hayashi, S., Y. Tanaka, and E. Kodama, "A New Manufacturing Control System using Mahalanobis Distance for Maximizing Productivity", *IEEE Transactions*, 59-62, 2001.

[9] Al-Otum, H.M., "Morphological Operators for Color Image Processing Based on Mahalanobis Distance Measure," *Optical Engineering*, 42 (9), 2595-2606, September 2003.

[10] Asada, M., "Wafer Yield Prediction by the Mahalanobis-Taguchi System", *IIE Transactions*, 25-28, 2001.

[11] Garcia-Lagos, F., Joya, G., Marin, F.J., and F. Sandoval, "Modular Power System Topology Assessment Using Gaussian Potential Functions," *IEE Proceedings-Generation Transmission and Distribution*, 150 (5), 635-640, September 2003.

[12] Wu, Y., "Pattern Recognition using Mahalanobis Distance", *TPD Symposium*, 1-14, 1996.

[13] Jugulum, R., and Monplaisir, L. "Comparison between Mahalanobis-Taguchi- System and Artificial Neural Networks", *Journal of Quality Engineering Society*, Vol. 10, No.1, pp.60-73, 2002.

[14] Fogel, L. J., Owens, A. J. and Walsh, M. J., "Artificial intelligence through simulated evolution," *Biophysics and Cybernetic Systems: Proceedings of the 2nd Cybernetic Sciences Symposium*, pages 131-155, 1965.

[15] Fogel, D. B., "Evolving behaviors in the iterated prisoner"s dilemma". *Evolutionary Computation*, Vol. 1, No. 1, 1993.

[16] Stanley, E. A., Ashlock, D. A. and Tesfatsion, L., "Iterated prisoner's dilemma with choice and refusal," *Artificial Life III*, vol. 17 of Santa Fe Institute Studies in the Sciences of Complexity, edited by Christopher Langton, Reading, 1994, Addison-Wesley, pp. 131-176.

[17] Chellapilla, K. and Czarnecki, D., "A preliminary investigation into evolving modular finite state machines," *Proceedings of the 1999 Congress on Evolutionary Computation*, 1999, pp. 1349-1356.

[18] Ashlock, D. A., Wittrock, A. and Wen, T., "Training finite state classifiers to improve PCR primer design," *Proceedings of the 2002 Congress on Evolutionary Computation*, IEEE Press, 2002, pp. 13-18.

[19] Ashlock, D. A., Bryden, K. M., Corns, S. M., Schnable, P. S., and Wen, T.J., "Training Finite State Classifiers to Improve PCR Primer Design," 10th Annual AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Albany, NY, 2004.

[20] Yadav, S. and Corns, S.M., "Improved PCR Design for Mouse DNA by Training Finite State Machines," *2010 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*, pages 249-254, 2010.

[21] Ercal, F., Chawla, A., Stoecker, W. V., Lee, H., and Moss, R. H., "Neural network Diagnosis of Malignant melanoma From Color Images," *IEEE Transactions on Biomedical Engineering*, vol. 41, No. 9, pages 837-845, September, 1994.

[22] Erkol, B., Moss, R. H., Stanley, R. J., Stoecker, W. V., and Hvatum, E., "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," Skin Research Technology, vol 11, pages 17-26, 2005.

[23] Swetter, S. M., "Malignant Melanoma," eMedicine (on-line source), http://emedicine.medscape.com/article/1100753-overview, last accessed 16 November, 2010