A MEMETIC ALGORITHM FOR GENERATING SPECTRAL INDICES FOR REMOTELY SENSED IMAGERY

Ivan Marc H. Escamos (1), Maria Art Antonette D. Clariño (1), Joseph Anthony C. Hermocilla (1), Jaderick P. Pabico (1), Cristino L. Tiburan, Jr. (1)

[1] University of the Philippines Los Baños, Los Baños, Laguna, 4031, Philippines
Email: ivanhescamos@gmail.com; mdclarino@up.edu.ph; jchermocilla@up.edu.ph; jppabico@up.edu.ph; cltiburan@up.edu.ph

KEY WORDS: Spectral Index, Memetic Computing, Search Algorithms, Feature Extraction

ABSTRACT: Spectral indices are formulas that integrate different wavelengths, or bands, of the electromagnetic spectrum to accentuate the abundance of various features of interest such as vegetation, burned areas, man-made, water, and geological features. Given its numerous applications, research on the development of spectral indices is a worthwhile undertaking. New spectral indices could be created as improvements over existing ones in aspects such as separability and sensitivity. New spectral indices can be developed for other features not yet covered by any spectral index.

The problem of developing spectral indices was posed as a search problem. The search space consists of all possible spectral indices. A spectral index was viewed as a combination of various spectral bands, mathematical operators, and numerical coefficients. Therefore, search algorithms could be used to generate spectral indices.

For this study, the memetic algorithm was utilized. Specifically, a memetic variant of genetic programming was developed. This was achieved by augmenting the genetic programming algorithm with the simulated annealing algorithm.

The developed algorithm was used to generate spectral indices for vegetation and built-up areas. Training points derived from Landsat 8 imagery was used as the input. The quality of the generated spectral indices was measured using two metrics: Silhouette Score and the Jeffries-Matusita distance. Indices for vegetation and built-up areas were developed: namely Memetic Genetic Programming Vegetation Index (MGPVI) and Memetic Genetic Programming Built-up Index (MGPBI). MGPVI was compared with other vegetation indices and MGPBI was compared with other built-up indices. Both generated indices outperformed their competing indices in terms of the earlier mentioned metrics.

1. INTRODUCTION

An index is a metric that describes complex phenomena using a limited number of parameters. (Bouzekri et al., 2015) In remote sensing, a spectral index is a formula that integrates different wavelengths, or bands, of the electromagnetic spectrum to accentuate the abundance of various features of interest such as vegetation, burned areas, man-made, water, and geological features. (Harris Geospatial Solutions, 2018) Spectral indices were also used by researchers as tools in aiding land cover classification in remotely sensed imagery. (Jabloun et al., 2009; Bhatt et al., 2018)

One of the most popular spectral indices is the vegetation index Normalized Difference Vegetation Index (NDVI) which was developed by Rouse. (Rouse et al., 1974). This index highlights areas of healthy, green vegetation. NDVI utilizes the near-infrared band (0.6 µm) and the visible red band (0.9 µm) and is given by the formula:

$$\text{NDVI} = \frac{\text{NIR - RED}}{\text{NIR + RED}} \tag{1}$$

where NIR represents the Near-Infrared band, and RED represents the red band. The RED band is used since healthy vegetation absorbs waves of this frequency well. On the other hand, the NIR band is used since healthy vegetation reflects this frequency well. (Perez and Comiso, 2015)

Given the numerous applications of spectral indices, research on this area is a worthwhile undertaking. New spectral indices could be created as improvements over existing ones in aspects such as separability and sensitivity. Spectral indices can be developed for other features not yet covered by any spectral index. Also, spectral indices could also be devised as feature extraction tools to improve the results of land cover classification

The problem of finding spectral indices can be posed as a search problem. The search space consists of all possible spectral indices. A spectral index can be viewed as a combination of various spectral bands, mathematical operators, and numerical coefficients. There could be better spectral indices waiting to be uncovered in the depths and breadths of the search space; convoluted formulas and unexpected combinations could be found. Thus, search algorithms and heuristics could be utilized to generate spectral indices.

Evolutionary Algorithms are search and optimization methods that were inspired by Darwinian Evolution. In EAs, the essence of natural selection is imitated. Solutions are represented as chromosomes, and these chromosomes are made to evolve by operations such as selection, crossover, and mutation. It has been established that pure evolutionary algorithms are not effective in complex combinatorial spaces, and hybridization with other approaches can highly improve the efficiency of search. The combination of EAs with local search operators was dubbed as Memetic Algorithms. These methods were based on adaptation models in natural systems that fuse the evolutionary adaptation of a population with the individual learnings of its members. The etymology of the name Memetic Algorithm was Richard Dawkins' concept of a meme, which represents "a unit of cultural evolution that can exhibit local refinement." The memetic model captures the plasticity of individuals that EAs cannot. (Krasnogor and Smith, 2005)

The main contribution of this study is to design, implement, and evaluate the utilization of memetic algorithms in generating spectral indices for remotely sensed imagery. The specific contributions of this study are as follows:

1. To identify various metrics to evaluate the quality of a spectral index
2. To utilize the developed algorithm for crafting spectral indices for various applications such as vegetation and urban sprawl monitoring; and
3. To compare the generated spectral indices with existing ones

2. REVIEW OF RELATED LITERATURE

Several literatures have reported on the use of Genetic Programming in the derivation of spectral indices. Chion, et al. authored A Genetic-Programming-Based Method for Hyperspectral Data Information Extraction: Agricultural Applications. (Chion, et al., 2008) In this study, the researchers introduced a method called genetic programming-spectral vegetation index or GP-SVI. Using Genetic Programming, a descriptive regression model is evolved, wherein ground truth observations of nitrogen content were correlated into a vegetation index. The fitness function for the algorithm consists of correlation strength with the ground truth data and length was penalized, meaning shorter indices were rewarded and longer indices were penalized. Two spectral indices for estimating nitrogen content were developed.

In a work by Puente, et al., the researchers created vegetation indices using Genetic Programming to automatically detect the sum of healthy, dry, and dead vegetation in the context of soil erosion assessment. The Revised Universal Soil Loss Equation or RUSLE is an erosion model defined by the equation A = R*K*L*S*C*P where R is the rainfall-runoff factor, K represents the influence of soil properties on soil loss during storm events, L represents the slope length factor, S is the slope steepness factor, C provides the ground cover factor, and P describes the conservation support practice factor. The C factor is one of the most vital factors and it can be derived using vegetation indices. However, the results tend to be incomplete since most vegetation indices are only sensitive to healthy vegetation, whilst senescent vegetation also helps in reducing erosion. The authors devised a novel approach where GP was used to calculate the C factor by creating spectral indices correlated with field and satellite information. (Puente, et al., 2011)

In 2017, Albarracin authored his Master's Thesis entitled "Genetic-Programming-Based Spectral Indices for Remote Sensing Image Classification." (Albarracin, 2017) In the study, the author developed an image classification algorithm based on spectral indices derived using Genetic Programming. The derivation of spectral indices was similar to the previous papers in this section; but instead of using correlation with ground measurements as the fitness function, a metric called the Silhouette Score was used. The spectral indices were used as a means of feature extraction and dimensionality reduction.

3. MATERIALS AND METHODS

3.1 Workflow for Generating Spectral Indices

Figure 1 illustrates the workflow which begins with a multiband remotely-sensed image. From the image, training points will be extracted. The training points will be the input to the memetic genetic programming algorithm. The algorithm will then output a spectral index which would be subjected into performance evaluation using various metrics described in the succeeding sections.

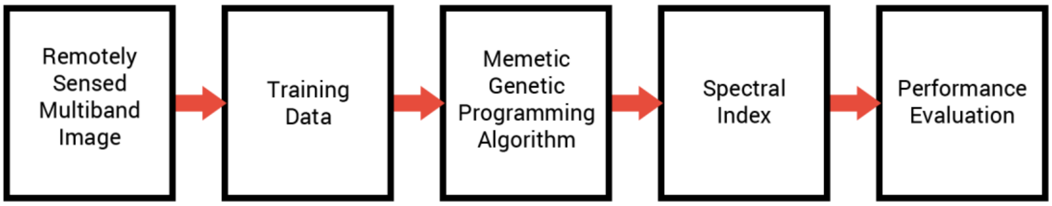

Figure 1. Workflow for Generating Spectral Indices

3.2 Metrics for the Comparison of Spectral Indices

3.2.1 Silhouette Score

The problem of creating spectral indices could be linked to clustering. Given a multiband image with n bands, calculating a spectral index merges the n-dimensional data into one dimension. In the multiband image, given a distribution of points of different cover classes, each class could be interpreted as a cluster. Points of the same class should have similar calculated spectral index values and could be grouped into a cluster. Therefore, in the design of a spectral index, index values of the same class should be as similar as possible, and index values of different classes must be as different as possible.

In the seminal paper by Rousseeuw, he introduced a metric called the Silhouette Score (SS). (Rousseeuw, 1987) The silhouette is based on the comparison of every cluster's tightness and separation. It quantifies how well an object is assigned to its cluster, and with respect to other clusters. The average dissimilarity of an object x to a cluster Ci is the mean of the distances of x to all objects that belong to Ci. Let a(x) be the average dissimilarity of x to its own cluster. b(x), on the other hand is the smallest average dissimilarity of x to the other clusters. b(x) shows the dissimilarity of x with another cluster that it is most similar to. The silhouette of an object x is given by:

$$s(x) = \frac{b(x) - a(x)}{max\big(a(x), b(x)\big)} \tag{2}$$

3.2.2 Jeffries-Matusita Distance

Another metric that can be used to quantify the quality of a spectral index is the Jeffries-Matusita distance (JMD). It is commonly used to measure the separability of land cover classes. JM distance is calculated as:

$$\mathrm{JM} = 2(1 - e^{-B}) \tag{3}$$

where B is the Bhattacharyya distance which is given by:

$$\mathrm{B} = \frac{1}{8}\mathrm{D}^2_{\mathrm{pool}} + \frac{1}{2} ln\left(\left|\frac{\Sigma_i + \Sigma_j}{2}\right| \Big/ \sqrt{|\Sigma_i| \cdot |\Sigma_j|}\right) \tag{4}$$

where, D^2_{pool} is the pooled Mahalanobis distance which is given by:

$$D^2_{pool} = (\mu_i - \mu_j)' \left(\frac{\Sigma_i + \Sigma_j}{2}\right)^{-1} (\mu_i - \mu_j) \tag{5}$$

where, i and j are classes, μ is the mean vector of index values, and Σ is the variance-covariance matrix. (Van Niel et al., 2005) The values of the JM distance range from 0 (indicates a low separability between classes) to 2 (indicates high separability). (Qiu et al., 2014)

3.3 Training Data

The workflow for generating spectral indices need points as training data. Given a remotely sensed image as a basemap, points were digitized in the ESRI shapefile format. Each point has a class label and the band reflectance values. In digitizing points, each class must be properly represented by a variety of examples for different instances of the class.

3.4 Memetic Genetic Programming Algorithm

3.4.1 Flow of the Algorithm

The memetic genetic programming algorithm that was implemented was based on the template of a memetic algorithm described in Cotta, et al. (Cotta, et al., 2016) An initial randomized population of spectral indices will be generated, and these spectral indices have an associated fitness value computed using a fitness function. The higher the fitness, the better the spectral index. The initial randomized population will then be improved using a local search operator. For a certain number of generations, the population will be evolved using the selection, crossover, mutation, and local search operations. The best spectral index, meaning the index with the highest fitness value, for all generations will be the output of the algorithm.

The following describes what happens in every generation. Using the selection operation, spectral indices will be chosen as parents based on their fitness values. The selected parents are recombined using the crossover operation and will result to two new "offspring" spectral indices. These offspring are improved using the local search operator. After improvement, the improved offspring will be further evolved using the mutation operation. After the mutation, the mutated offspring will be improved again using the local search operator. These improved offspring will now replace the current population. The population is considered stagnant if the best spectral index does not change after a certain number of generations. If the population is stagnant, the restart operation commences where the worse half of the population is replaced with a new set of spectral indices improved using the local search operator.

3.4.2 Genetic Representation

Spectral indices were represented as syntax trees. The root and internal nodes represent mathematical operations, specifically addition (+), subtraction (-) multiplication (*), and protected division (/). Instead of a regular division operation, a safe division operation was implemented to handle complications that could be brought up by dividing by zero, which is undefined. When the divisor is zero, the quotient will be set to one. On the other hand, the leaf nodes contain the composition of a selected band, a coefficient between zero and one, and a unary operator. The unary operators in the leaf node comprise of a negation operation (-), protected square root (sqrt), and protected base-10 logarithm (log). The selected band should be a valid band from the type of spectral image used. A safe square root operation was implemented to handle complications that could be caused by negative operands. This is done by always applying the absolute value operation on the operand. Also, a safe base-10 logarithm was implemented to handle zero and negative values by returning zero if the input is zero and getting the base-10 logarithm of the absolute value of the operand when the values are negative. The safe operators discussed were recommendations from a study by Koza. (Koza, 1994). Figure 2 shows the syntax tree representation of an index $\log_{10}B1$ * sqrt(B7) and the Normalized Difference Vegetation Index.

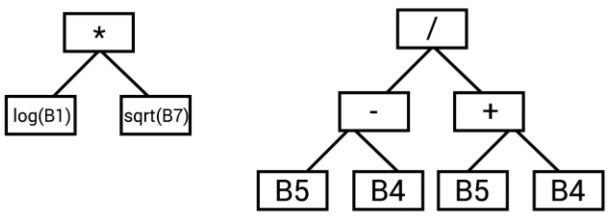

Figure 2. Syntax tree representation of an index $\log_{10}B1$ * sqrt(B7) and the Normalized Difference Vegetation Index

3.4.3 Fitness Function

The fitness function for the candidate spectral indices of the algorithm is given below:

$$F = a \cdot silhouetteScore + b \cdot JMDistance/2 + c \cdot selectedClass + d \cdot otherClasses + e \cdot inRange + f \cdot isClassMax$$
$$where\ a + b + c + d + e + f = 1.0 \tag{6}$$

F is a weighted function of two factors: quality and form. The quality factor measures the quality of the spectral index. For the quality factor, the silhouette score (silhouetteScore) and the Jeffries-Matusita Distance (JMDistance) of the spectral index were calculated using the training points. The form factors are boolean values that indicate some trait in the calculated values of the spectral index. It includes traits such as: the mean value of the calculated index values for the selected class must be greater than a value w (selectedClass), the mean value of the calculated index values for the other classes must be less than a value x (otherClasses), and both mean values must be within values y and z (inRange), and if the largest calculated index value belongs to the selected class (isClassMax).

The form factors, on the other hand, ensure that the index adheres to some value constraints. The form factors will help standardize the calculated values of the spectral index to make it easily interpretable

3.4.4 Selection

For the selection operation, roulette wheel selection was used. In roulette wheel selection, the probability of an individual to be selected is proportional to its fitness. In this method, the population is sorted based on their fitness (F) in descending order. The fitness values are then normalized in a range from 0 to 1. A random number between 0 and 1 is generated. The sorted list of individuals is traversed, accumulating the normalized fitness scores. When the accumulated value exceeds the randomized number, the individual is selected.

3.4.5 Crossover

The crossover operation involves the exchange of subtrees between two syntax trees. For the simplicity of the implementation, the crossover point is at the root node. Figure 3 shows an example of a crossover between two spectral indices.

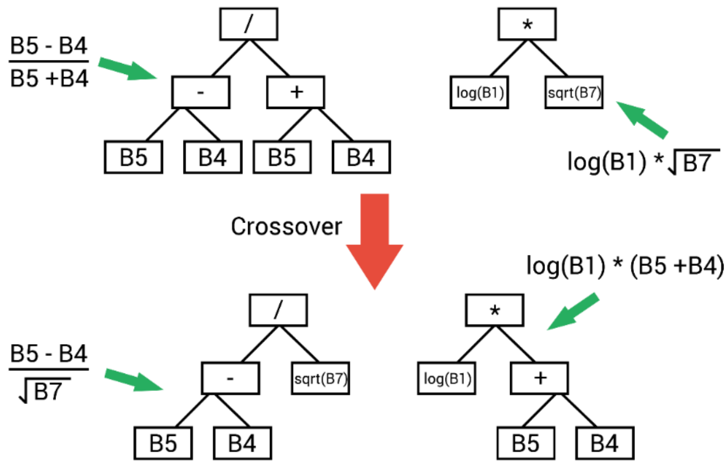

Figure 3. Sample crossover between two spectral indices

3.4.6 Mutation

Two cases of mutation were implemented: basic and expansion. The basic case includes traversing the syntax tree, and according to the mutation rate, change the content of a node. Expansion involves randomly generating a spectral index with a height of two and replacing a leaf node of the original syntax tree with the randomly generated one. Figure 4 shows examples of a mutation operation.

3.4.7 Local Search Operator: Simulated Annealing

Tampering metal, glass, or crystal by heating above its melting point, maintaining its temperature, and cooling it very slowly until it solidifies into a perfect structure is referred to as annealing. Emulating this process as a search heuristic is called simulated annealing. In the computer science analogy, the perfect crystalline structure represents the best solution, the physical material represents problem solutions, the energy of the state represents the fitness of a solution, and the temperature represents a control parameter. (Yang, 2010) Simulated Annealing is performed on a spectral index to improve it.

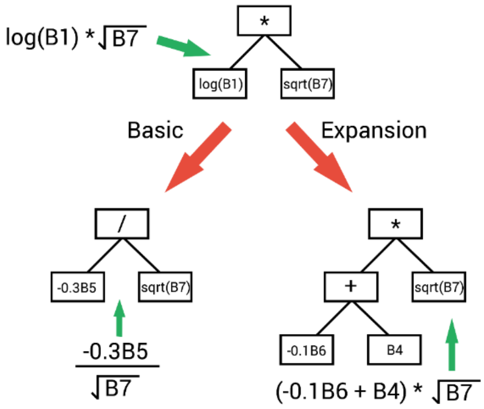

Figure 4. Sample mutation operation

The following briefly describes the algorithm. The current temperature is represented by T and is usually initialized with a high value. Given a certain number of iterations, the spectral index is attempted to be improved by performing modifications on the index. If the fitness improved, it replaces the current index. Else, the index is replaced with a probability of $e^{(\Delta fitness/T)}$, where Δ*fitness* is the difference of the fitness values before and after the attempted improvement, and T is the current temperature. After the series of attempted improvements, T is decreased by multiplying it with the factor (*1 - cooling*), where *cooling* is the cooling schedule typically a value under 1.0. The process stops until T_{min} or the minimum set temperature is reached. The best index during the entire process is the output of the algorithm.

3.5 Input Data

For the input multiband image, a Landsat 8 scene (Path 116, Row 51) taken August 1, 2016 was used. The image was pre-processed using ENVI software's radiometric calibration tool to convert the raw digital numbers to actual reflectance values.

Five classes were defined: (1) Vegetation, (2) Built-up, (3) Open/Bare Areas, (4) Sediment/Sand, and (5) Water. Five-hundred training points were collected from the imagery with 100 points per class. Figure 5 shows the input Landsat 8 scene with the training points overlain.

3.6 Validation Data

A new validation point dataset was created to remove the bias for the calculation of the index's silhouette score. The validation set consists of 500 points with 100 points per class, similar to the training data. The validation dataset was collected only from a subregion of the image. This was due to computational constraints of calculating the spectral index images. Figure 12 shows the subregion of the Landsat Image and the collected validation points.

3.7 Hyperparameters of the Algorithm

The Memetic Genetic Programming Algorithm was ran with the following hyperparameters: Number of Generations = 200, Population Size = 20, Crossover Rate = 0.90, Mutation Rate = 0.90, Stagnant If = 20 generations, a = 0.30, b= 0.30, c = 0.1000, d = 0.10, e = 0.10, f= 0.1, w = 0.50, x = 0.50, y = -1, and z = 1. The Simulated Annealing Algorithm was ran with Initial Temperature = 5000, T_{min} = 1, Cooling = 0.3, and Number of Iterations = 10. These hyperparameters were both used in generating both MGPVI and MGPBI.

4. RESULTS AND DISCUSSION

Using the developed Memetic Genetic Programming Algorithm for generating spectral indices, two prototype spectral indices have been generated using Landsat Imagery. These indices are the Memetic Genetic Programming Vegetation Index (MGPVI), and Memetic Genetic Programming Built-up Index (MGPBI). Sample images for these indices are shown in Figure 7.

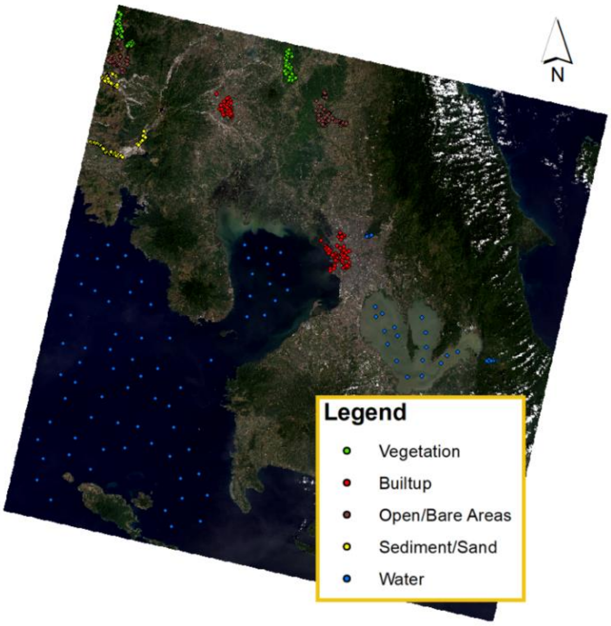

Figure 5. Input Landsat 8 Scene and Training Points

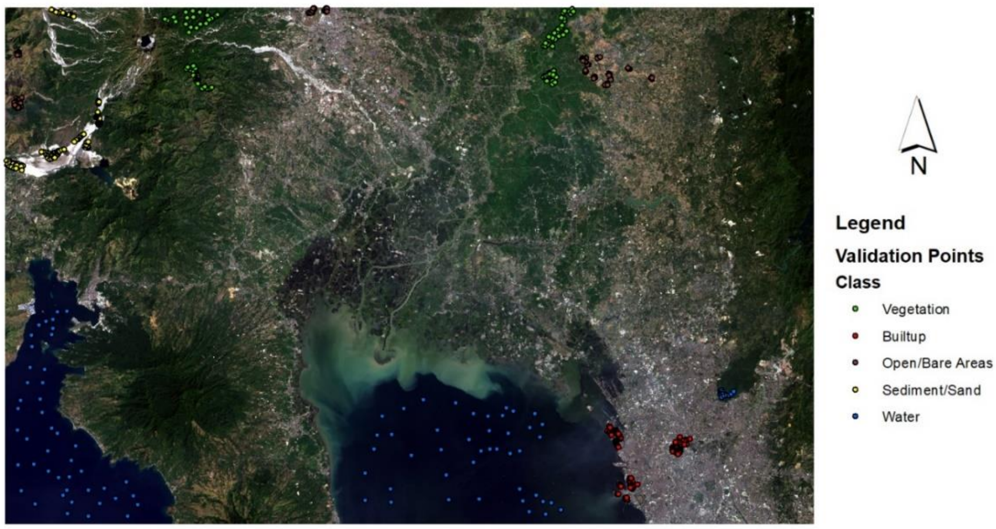

Figure 6. Landsat 8 Scene Subregion and Validation Points

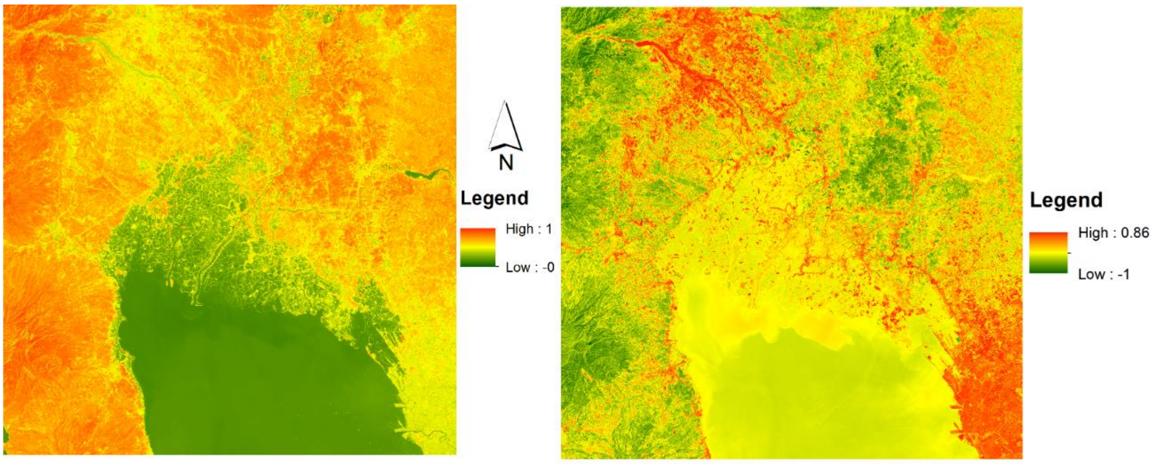

Figure 7. Sample outputs for MGPVI (Left) and MGPBI (Right)

4.1 Memetic Genetic Programming Vegetation Index

MGPVI was defined as:

$$MGPVI = \frac{0.35B5 \cdot \log_{10} 0.83B3 \cdot \sqrt{0.75B2}}{0.88B1 \cdot \left(\sqrt{0.7B6} - \sqrt{B5}\right)} \tag{7}$$

MGPVI was compared to other vegetation indices: Simple Ratio (SR) (Birth and McVey, 1968), Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI) (Huete, et al., 1980), and Global Environment Monitoring Index (GEMI) (Pinty and Verstraete, 1992). These indices are given by the following equations:

$$\mathrm{SR} = \frac{\mathrm{NIR}}{R} \tag{8}$$

$$\mathrm{SAVI} = \frac{(NIR - R)}{(NIR + R + L)}(1 + L) \tag{9}$$

$$\mathrm{GEMI} = \eta(1 - 0.25\eta) - (R - 0.125)/(1 - R)$$

$$\eta = \frac{2(NIR^2 - R^2) + 1.5NIR + 0.5R}{NIR + R + 0.5} \tag{10}$$

Using the validation data, the metrics of the vegetation indices were calculated and given by the table below:

Table 1. Calculated Metrics for Vegetation Indices

Vegetation Index	Silhouette Score	Jeffries-Matusita Distance
MGPVI	**0.70**	**2.00**
SR	0.58	1.81
NDVI	0.64	2.00
SAVI	0.54	2.00
GEMI	0.61	2.00

It can be observed from Table 1 that MGPVI garnered the highest score from all the indices.

4.2 Memetic Genetic Programming Built-up Index

MGPBI was defined as:

$$MGPBI = \sqrt{0.39B7} + \sqrt{0.94B2} - \frac{-0.94B5}{\log_{10} 0.57B6 + \log_{10} 0.44B5}$$

MGPBI was compared to other built-up indices: Normalized Difference Built-up Ratio (NDBI) (Zha, et al., 2003), Urban Index (UI) (Kawamura, et al., 1996), and Index-based Built-up Index (IBI) (Xu, 2008). These indices are given by the following equations:

$$\text{NDBI} = \frac{(SWIR\ -\ NIR)}{(SWIR\ +NIR)}$$

(11)

$$UI\ =\frac{SWIR\_2\ -\ NIR}{SWIR\_2\ +\ NIR}$$

(12)

$$IBI\ =\frac{\frac{2SWIR}{SWIR+NIR}-\left(\frac{NIR}{NIR\ +\ RED}+\frac{GREEN}{GREEN+SWIR}\right)}{\frac{2SWIR}{SWIR+NIR}+\left(\frac{NIR}{NIR\ +\ RED}+\frac{GREEN}{GREEN+SWIR}\right)}$$

(13)

Using the validation data, the metrics of the built-up indices were calculated and given by the table below:

Table 2. Calculated Metrics for Built-up Indices

Built-up Index	Silhouette Score	Jeffries-Matusita Distance
MGPBI	**0.56**	**1.99**
NDBI	0.12	1.99
UI	0.15	1.98
IBI	0.15	1.99

It can be also observed from Table 2 that MGPBI garnered the highest score among the built-up indices.

5. CONCLUSION AND FUTURE WORK

Given the many applications of spectral indices, a Memetic Genetic Programming Algorithm was developed to search for spectral indices. The flow of the algorithm along with its operators were discussed. Two measures of quantifying the quality of the spectral index were discussed: Silhouette Score and the Jeffries-Matusita Distance. Two indices were developed using the algorithm: Memetic Genetic Programming Vegetation Index and Memetic Genetic Programming Built-up Index. These generated indices were compared to other indices using the measures of index quality described in the paper. It was observed that both indices perform better than the others.

The developed algorithm still has room for improvement. Many hyperparameter value combinations have not been tried out. Spectral indices for features currently without indices could also be discovered using the algorithm. In a future study, the derived indices will be used as a form of feature extraction to improve the performance of machine learning classifiers in land cover classification

6. ACKNOWLEDGMENT

This work is funded by Department of Science and Technology-Science Education Institute. The author would also like to acknowledge his advisory committee for the support in his research endeavors.

7. REFERENCES

Albarracin, F.H., 2017. Genetic-programming-based spectral indices for remote sensing image classification. Master’s thesis. University of Campinas

Birth, G. S., and G. R. McVey. 1968. Measuring the Color of Growing Turf with a Reflectance Spectrophotometer. Agronomy. J. 60, pp. 640-643.

Bhatt, A., Ghosh, S.K., and Kumar, A., 2018. Spectral indices based object oriented classification for change detection using satellite data. International Journal of System Assurance Engineering and Management, 9(1), pp. 33–42

Bouzekri, S., Lasbet, A.A., and Lachehab, A., 2015. A new spectral index for extraction of built-up area using landsat-8 data. Journal of the Indian Society of Remote Sensing, 43(4), pp. 867–873

CHION, C., LANDRY, J., and COSTA, L.D. 2008. A genetic-programming-based method for hyperspectral data information extraction: Agricultural applications. IEEE Transactions on Geoscience and Remote Sensing, 46(8), pp. 2446–2457

Cotta, C., Mathieson, L., & Moscato, P. (2018). Memetic algorithms. Handbook of Heuristics, pp. 607-638.

Harris Geospatial Solutions, 2018. Spectral Indices, Retrieved July 2, 2018 from http://www.harrisg eospatial.com/ docs/SpectralIndices.html.

Huete, A., 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment, 25(3), pp. 295– 309

Jabloun, M., Mihai, C., Vanhamel, I., Geerinck, T., and Sahli, H., 2009. Multispectral data classification based on spectral indices and cascaded fuzzy c-mean classifiers. 2009 IEEE International Geoscience and Remote Sensing Symposium, volume 3, pages III–821–III–824

Kawamura, M., Jayamana. S., Tsujiko, Y., 1996. Relation between social and environmental conditions in Colombo Sri Lanka and the urban index estimated by satellite remote sensing data. Int. Arch. Photogramm. Remote Sens., 31 (Part B7), pp. 321–326.

Koza, J.R., 1994. Genetic programming as a means for programming computers by natural selection. Statistics and computing, 4(2), pp. 87–112

Krasnogor, N. and Smith, J., 2005. A tutorial for competent memetic algorithms: model, taxonomy, and design issues. IEEE Transactions on Evolutionary Computation, 9(5), pp. 474–488

Perez, G., and Comiso, J., 2015. Monitoring Philippine vegetation using satellite ndvi and evi data. Journal of the Philippine Geosciences and Remote Sensing Society, 1(1), pp. 27–40

Pinty, B., and Verstraete M.M., 1992. Gemi: a non-linear index to monitor global vegetation from satellites. Vegetatio, 101(1), pp. 15–20

Puente, C., Olague, G., Smith, S.V., Bullock, S.H., Hinojosa-Corona, A., Gonzales-Botello, M.A., 2011. A genetic programming approach to estimate vegetation cover in the context of soil erosion assessment

Qiu, B., Fan, Z., Zhong, M., Tang, Z., Chen, C. 2014., A new approach for crop identification with wavelet variance and JM distance. Environmental monitoring and assessment, 186.

Rouse Jr, J., Haas, R., Schell, J., and Deering, D., 1974. Monitoring vegetation systems in the great plains with erts

Rousseeuw, P.J., 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, pp. 53 – 65

USGS, 2017. What is Landsat and when did it begin?. Retrieved July 2, 2018, from https://landsat.usgs.gov/ what-landsat-and-when-did-it-begin/.

Van Niel, T.G., McVicar, T.R., Datt, B., 2005. On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification. Remote Sensing of Environment. 98(4), pp 468-480

Xu, H., 2008. A new index for delineating built-up land features in satellite imagery. Int. J. Remote Sens. 29, pp. 4269–4276

Yang, X.S., 2010. Nature-inspired metaheuristic algorithms, Luniver press

Zha, Y., Gao, J., and Ni, S., 2003. Use of normalized difference built-up index in automatically mapping urban areas from tm imagery. International Journal of Remote Sensing, 24(3), pp. 583– 594