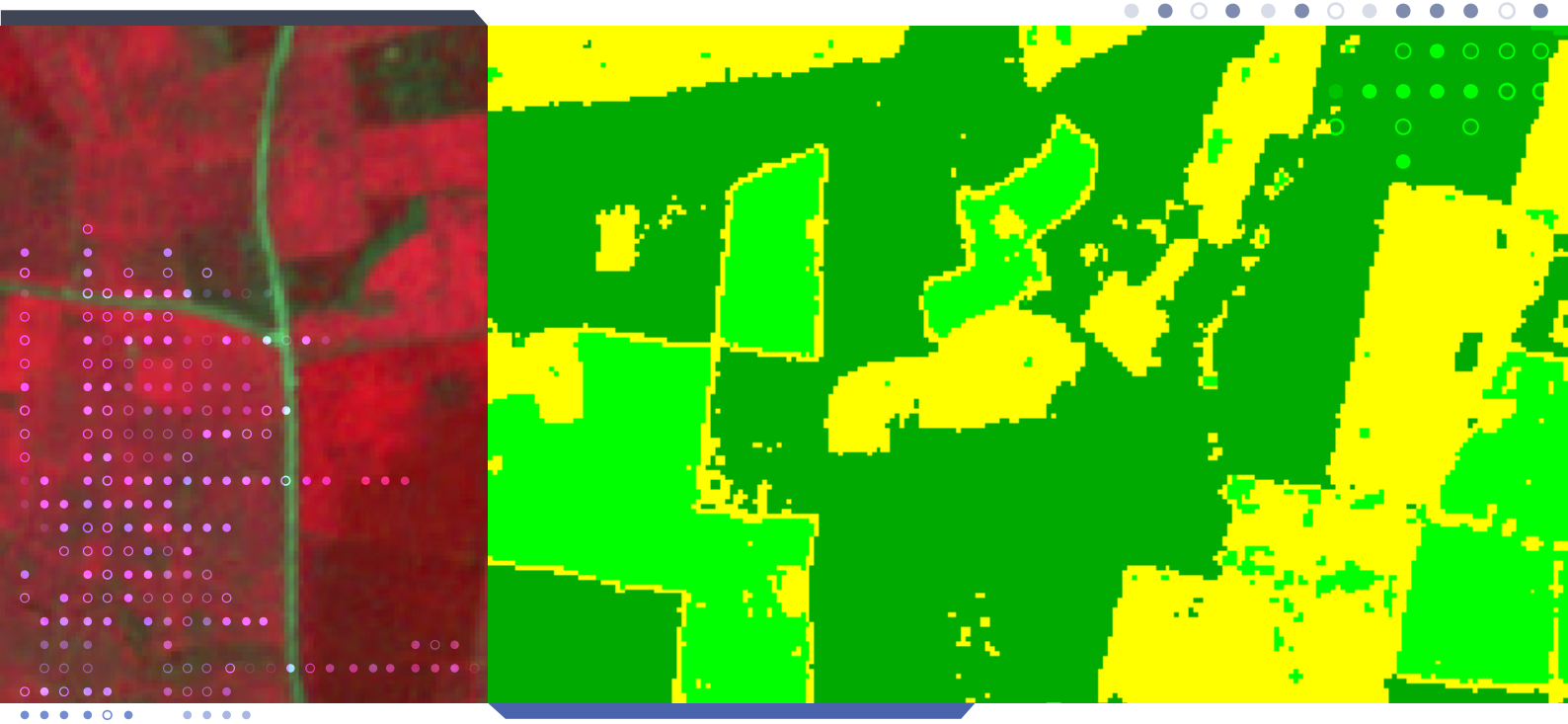
SCANPOINT GEOMATICS LIMITED

Machine Learning based Classification in IGiS

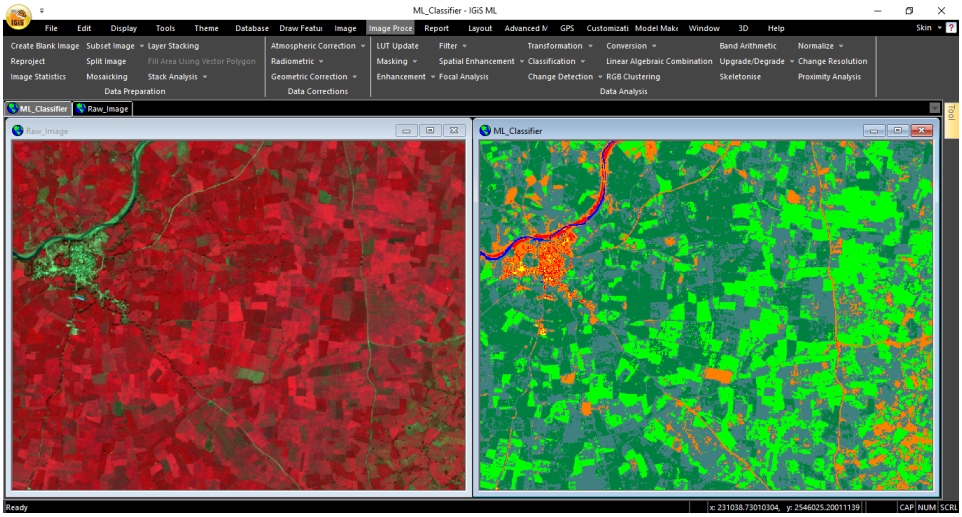


White Paper 2023

Exploring and understanding the capabilities of machine learning algorithms, such as Random Forest and Support Vector Classification, for precise satellite image classification has become crucial.

IGiS provides a robust solution for creating classified image with high level of accuracy using Random Forest Classifier (RFC) and Support Vector Machine Classifier (SVC). These classifiers have capabilities to classify the satellite images based on the training set and various input parameters provided by the users.

RFC's ability to handle multi-class classification tasks, and its robustness against overfitting by using training split parameter makes it suitable for highly accurate classification. On the other hand, SVC has the ability to handle non-linear classification tasks and its resilience against high-dimensional data makes it applicable for accurately classifying land cover patterns in satellite images.



IGiS provides the facilities to alter various parameters that enables the classifier for categorizing the data into probable classes. It can also provide facilities to generate area statistics, confusion matrix, etc. User can achieve this by using the following:

User-defined Hyper-parameters

Hyper-parameter Tuning

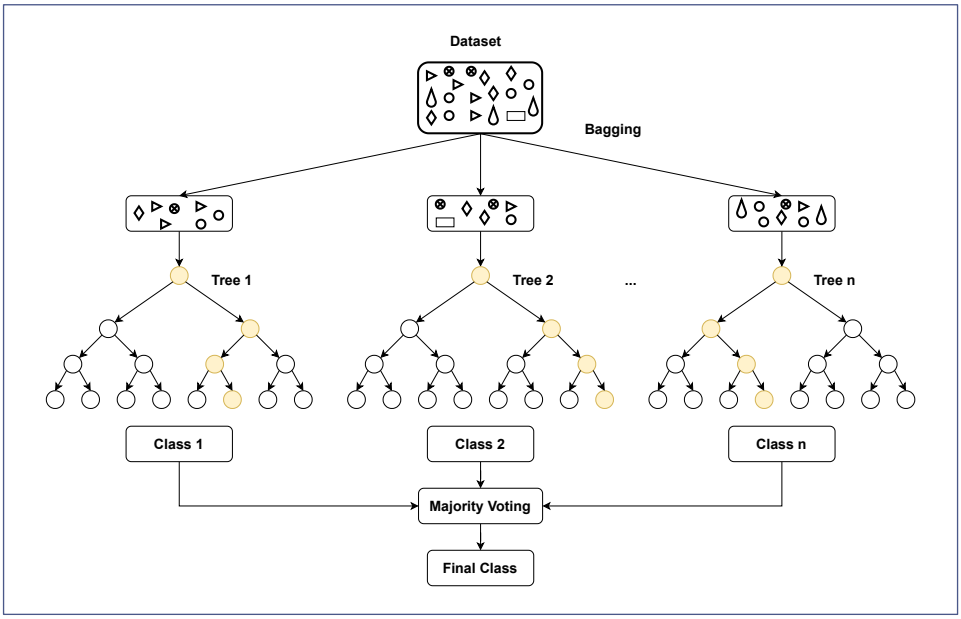
Training and Testing Split

Accuracy Reports

User-defined Hyper-parameters

IGiS provides the options to tune the hyper-parameters for training a model suitable for the classification of satellite image provided. Both RFC and SVC, have tuning parameters that can be adjusted to optimize their performance.

Random Forest Classification:



The hyper-parameters that can be tuned for RFC are as follows:

Number of Trees:

This parameter determines the number of decision trees in the random forest ensemble.

Maximum Depth of Trees:

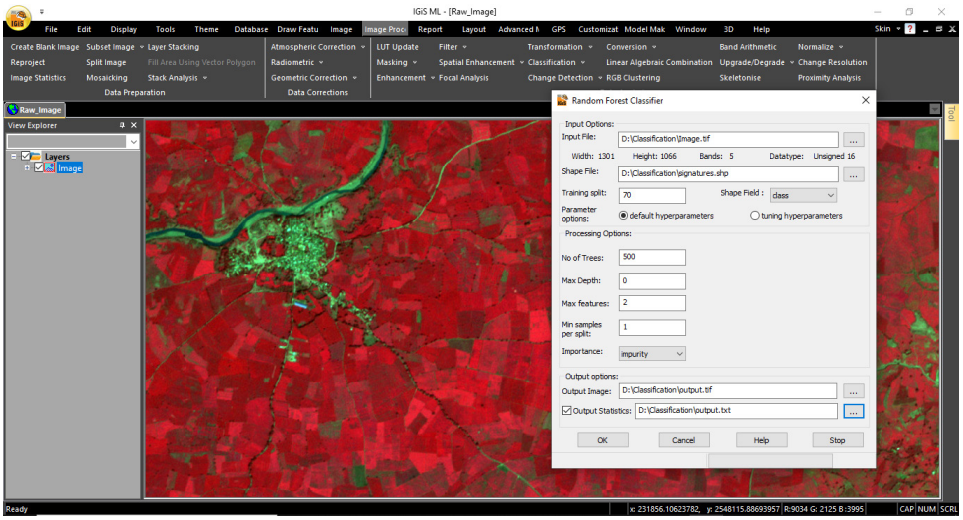
It controls the maximum depth of each decision tree in the random forest.

Minimum Samples for Split:

This parameter sets the minimum number of samples required to split a node in a decision tree.

Feature Subset Size:

It determines the number of features to consider when looking for the best split at each node.



Support Vector Classification:

The hyper-parameters that can be tuned for SVC are as follows:

Regularization Parameter (C):

This parameter controls the trade-off between maximizing the margin and minimizing the classification error.

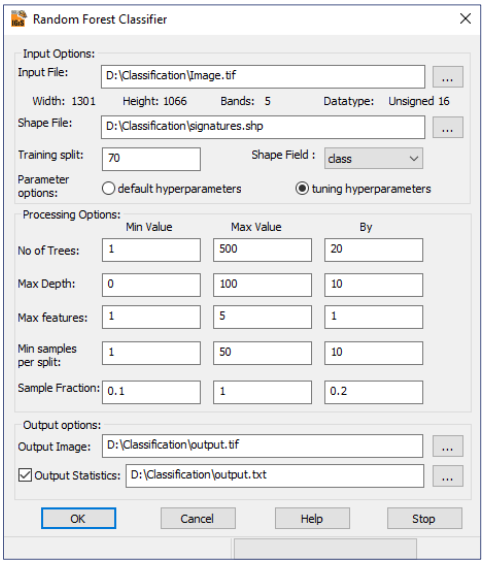
Kernel Type and Parameters:

SVC allows different kernel functions (e.g., linear, polynomial, radial basis function (RBF), and sigmoid) to transform the data into a higher-dimensional space. Each kernel function may have its specific parameters (e.g., polynomial degree, RBF kernel width).

Hyper-Parameter Tuning

Hyper-parameter tuning can be performed by employing techniques like grid search or random search that explore different combinations of parameter values and evaluate the model's performance using cross-validation.

IGiS technology provides inbuilt function for the selection of optimum hyper-parameters using grid search to achieve the best possible accuracy. This aids in the classification process and saves user's time for determination of the optimal hyper-parameters.



Training & Testing Split

It is crucial to split the available data into training and testing sets, when applying machine learning algorithms like RFC and SVC for satellite image classification. This split allows for training the model on a subset of the data and for independent evaluation on unseen data.

IGiS has the provision for the creation of training and testing dataset by using the training split parameter. Additionally, IGiS provides separate accuracy assessment for both datasets, so that the user understands how the model has performed on unseen data.

Accuracy Report

IGiS has inbuilt functions which can generate accuracy report for the classification performed by the generated machine learning model. The accuracy report includes the confusion matrix for both the training and testing datasets along with the overall accuracy, kappa coefficient, etc. These insights enable the user to assess the model's reliability and identify misclassified classes to make informed decisions for the executed classification.

```
Testing
Confusion Matrix and Statistics

          Reference
Prediction 1 2 3 4 5 6 7
         1 3 0 0 0 0 0 0
         2 0 3 0 0 0 0 0
         3 0 0 3 0 0 0 0
         4 0 0 0 2 0 0 0
         5 0 0 0 0 3 0 0
         6 0 0 0 0 0 2 0
         7 0 0 0 1 0 1 3

Overall Statistics

               Accuracy : 0.9048
                 95% CI : (0.6962, 0.9883)
    No Information Rate : 0.1429
    P-Value [Acc > NIR] : 1.376e-14

                  Kappa : 0.8889

 Mcnemar's Test P-Value : NA

Statistics by Class:

                     Class: 1 Class: 2 Class: 3 Class: 4 Class: 5 Class: 6  Class: 7
Sensitivity            1.0000   1.0000   1.0000  0.66667   1.0000  0.66667  1.0000
Specificity            1.0000   1.0000   1.0000  1.00000   1.0000  1.00000  0.8889
Pos Pred Value         1.0000   1.0000   1.0000  1.00000   1.0000  1.00000  0.6000
Neg Pred Value         1.0000   1.0000   1.0000  0.94737   1.0000  0.94737  1.0000
Prevalence             0.1429   0.1429   0.1429  0.14286   0.1429  0.14286  0.1429
Detection Rate         0.1429   0.1429   0.1429  0.09524   0.1429  0.09524  0.1429
Detection Prevalence   0.1429   0.1429   0.1429  0.09524   0.1429  0.09524  0.2381
Balanced Accuracy      1.0000   1.0000   1.0000  0.83333   1.0000  0.83333  0.9444
```

SCANPOINT GEOMATICS LIMITED

CONCLUSION

Overall, the incorporation of machine learning algorithms empowers researchers and professionals to attain heightened accuracy levels while minimizing the need for user intervention. By leveraging IGiS technology, organizations can effectively adopt AI-ML based classification techniques and conduct diverse analyses to generate the desired outcomes.

ABOUT

Scanpoint Geomatics Limited

Scanpoint Geomatics Ltd. is the leader in the Indian Geomatics Industry. We pioneer the nation's geospatial domain through IGiS. An indigenous technology that brings GIS, Image Processing, and Photogrammetry together on the same platform under the Make in India Initiative. We are proud of our partnership with the Indian Space Research Organisation (ISRO). With an innovative approach and over two decades of rigorous research and development, the duo developed the IGiS platform. Backed by ISRO's domain expertise, we aim to push forth innovation and uplift the global geospatial domain.

10th Floor, Swati Clover, Shilaj Circle, Sardar Patel Ring Road,
Thaltej, Ahmedabad, Gujarat-380054, India.

079 4602 3912 | info@sgligis.com
www.sgligis.com