International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 4, August 2024, pp. 3127~3136
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i4.pp3127-3136

Holdout based blending approaches for improved satellite image classification

Suresh Kumar Musali[1], Rajeshwari Janthakal[1], Nuvvusetty Rajasekhar[2]
[1]Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, Affiliated to Visvesvaraya Technological University, Belagavi, India
[2]Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, Affiliated to Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Oct 8, 2023
Revised Jan 29, 2024
Accepted Mar 13, 2024

Keywords:

Blending
Holdout
Machine learning
Remote sensing
Satellite image classification

ABSTRACT

An essential component of remote sensing, image analysis, and pattern recognition is image categorization. The classification of land use using remotely sensed data creates a map-like representation as the final form of the investigation. With its ability to effectively categorize satellite images, machine learning (ML) algorithms have gained significant traction in a number of fields, including land-use planning, disaster response, and natural resource management. Ensemble learning is also a widely used technique for enhancing the precision of satellite image categorization, which combines multiple models to get more precise predictions. Holdout is an ensemble technique, where multiple ML algorithms are used for training on the same dataset. The primary goal of this study is to create a holdout model for classifying satellite images. Initially, this study explores the usage of ML algorithms namely support vector machines (SVM), k-nearest neighbor (KNN), decision trees (DT), gradient boosting classifier (GBC), histogram-based GBC (HGBC), random forest classifier (RF), bagging classifier (BC), XGBoost classifier for classifying satellite images. Later, GBC, HGBC, RF, BC, and XGBoost are combined to build a stacking model. The bagging ensemble model outperforms all other methods and reaches an accuracy of 88.90%. Finally, blending models with holdout approach were developed and achieved accuracy of 93.70%, 94.14%, and 93.87% which outperformed all previous algorithms.

Corresponding Author:

Suresh Kumar Musali
Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, Affiliated to Visvesvaraya Technological University
Shavige Malleshwara Hills, Kumaraswamy Layout, Bangalore-560111, Karnataka, India
Email: sureshkumar-ise@dayanandasagar.edu

1. INTRODUCTION

The subject area of satellite image processing is concerned with the evaluation and interpretation of information gathered by Earth-orbiting remote sensing satellites. These satellites utilize a variety of sensors, including cameras and radar, to take pictures of the Earth's surface, which may provide important details about everything from weather patterns to land use trends. Dealing with the enormous volumes of data that these satellites can gather is one of the primary problems of satellite image processing. Advanced computer algorithms are employed to interpret and analyze raw satellite images since the data is frequently too big and complicated for people to process manually. Image enhancement, categorization, and feature extraction are some of the crucial methods utilized in satellite image processing. The process of feature extraction is defined as locating particular elements, such as roads and houses, inside a picture. Numerous industries, including environmental monitoring, agriculture, and urban planning, use satellite image processing. For instance, agriculture yields, urban growth, and deforestation rates can be tracked using satellite images. Therefore, Satellite image processing is crucial for comprehending and handling our planet's resources as technology and computer power develops. The existing methods for classifying images captured by satellites can be improved in terms of performance and simplicity. This research suggests a potentially quite successful deep learning-based method for classifying and labeling satellite images.

Usage of remote sensing image-classification benchmark-128 (RSI-CB-128), a crowd-sourced geographic dataset comprising 36,707 images split across 45 classes, to train our designs. A few fully connected layers, such as global average pooling, and dense layers with activation functions, such as softmax and rectified linear unit (ReLU), were included, along with several deep learning approaches, including transfer learning [1]. The performance of supervised scene classification is limited by the lack of labeled remote sensing images as compared to the field of natural images, and unsupervised methods are not as suitable for practical applications. Thus, this paper presents a generative adversarial network-based semi-supervised remote sensing image scene classification technique [2].

These days, land usage land cover (LULC) can be obtained faster relatively because of the quick advancements in machine learning, geographic information systems, and remote sensing. To categorize land use, the data from satellite images is processed. This approach is more advantageous than traditional field surveys in the real area in a number of ways, including savings on time and money [3]. To increase classification accuracy, satellite images must be effectively classified. This research proposes the use of optimal feature selection in conjunction with hierarchical framework and ensemble learning (HFEL) for accurate satellite image recognition. The hierarchical framework image is used to extract relevant features using three distinct types of convolutional neural networks (CNN): AlexNet, LeNet-5, and a residual network (ResNet) [4].

This work uses an ensemble model made up of the following components to provide a novel staggered training approach: i) A high-accuracy, resource-intensive vision transformer and ii) a low-count convolutional neural network with quick training, but lower accuracy. An accurate base model that is scalable is offered by the vision transformer. The ensemble model of a CNN rapidly absorbs fresh data [5].

The state-of-the-art automated satellite image classification techniques, such as nearest neighbors, naive Bayes (NB), support vector machine (SVM), discriminant analysis, random forests (RF), decision trees (DT), semi-supervised, CNN models, deep CNN, and hybrid approaches, are the main focus of this paper after presenting the conventional supervised techniques [6]. One of the more alluring solutions is optical remote sensing since it provides vegetation indices and some data are freely available. Sentinel-2A in particular provides some vegetation indices calculated to assess vegetation status. It has a multispectral sensor (MSI) with blue, green, red, and near-infrared-1 bands at 10 m; red edge 1 to 3, near-infrared-2, and shortwave infrared 1 and 2 at 20 m; and three atmospheric bands (band 1, band 9, and band 10) at 60 m [7].

Because it makes it possible to compute vegetation indices, which are helpful for evaluating the condition of vegetation, optical remote sensing is one of the most alluring methods for creating crop cover maps. Numerous vegetation indices are available from the Sentinel-2A multispectral instrument (MSI), a multispectral sensor with 13 bands that covers the visible, near infrared, and short-wave infrared (SWIR) wavelength areas [8]. Preeminent methods for predicting agricultural land usage primarily depend on data that is sensed locally, including farmer questionnaires conducted during field visits and rainfall observations. Although locally sensed data offer rich information, they are very expensive to gather, loud, and challenging to scale. A potential remedy is provided by the combination of contemporary machine learning techniques with remote sensing and satellite imagery data, which are inexpensive and widely available resources [9].

Deep neural networks (DNNs) have gained significant traction in the field of remote sensing recently and are used for a variety of applications. Nevertheless, more work needs to be done to improve the DNNs' robustness and generalizability in order to achieve higher accuracy for a wider range of sensing geometries and categories [10]. Because machine learning algorithms can detect nonlinear connections, they are being used more and more in remote sensing applications. Many real-world applications have incorporated ensemble algorithms to increase prediction accuracy. We present a synopsis of three popular ensemble methods: stacking, boosting, and bagging [11].

Among the many uses for satellite images is land cover and crop (LCC) mapping, which requires accurate classification. Due to its ability to mix and integrate numerous classifiers, ensemble classifiers have demonstrated extraordinary performance in satellite image categorization in recent years. This study introduced per-pixel accuracy-based ensemble of extreme learning machine (PAELM), a revolutionary approach for classifying satellite images [12].

Satellite imagery is essential for environmental monitoring, law enforcement, and disaster assistance. Some users require assistance from humans to manually identify the buildings and objects in the pictures. Automation is essential because of the large locations that must be explored and the limited number of analysts available. To address the issue, however, conventional methods for item identification and classification need to become more accurate. For automating these tasks, a class of machine learning methods known as "deep learning" has showed promise [13].

The Brovey transform, principal component analysis, and CNN have been proposed as an efficient method of classifying satellite images. This method has been used to multispectral Landsat 8 operational land imager (OLI) satellite imageries. To distinguish between different land cover classifications, the descriptions are divided into five classes. The computation of the kappa coefficient and other accuracy systems of measurement are used to evaluate the predicted technique. The obtained data verify that the anticipated method produces results with a high degree of precision. A comparative analysis of the results using the more sophisticated procedures shows that the projected strategy outperforms the alternative methods. As a result, the proposed method can be effectively applied to address the challenges associated with land cover classification [14].

This study develops a system to categorize satellite images and extract data using image processing methods. Satellite images have been divided into usable and unused areas, and each of the classes has been further divided into four groups. Unused satellite images are separated into forest, river, desert, and beach areas, while usable images are further categorized into residential, commercial, industrial, and agricultural areas. This study concentrated on effective classification of satellite images [15].

For many practical applications, the classification of hyperspectral images (HSI) is widely used in the analysis of remotely sensed satellite pictures. This study combined the HybridSN and inception residual network architectures with a suggested architecture that was inspired by Inception. Better spectral-spatial learning features are made possible by the proposed architecture's 3D and 2D inception blocks. Three well-known public HSI datasets were used for the studies [16].

In many countries, spatiotemporal land use and land cover change (LULCC) is accelerated by population increase, resulting in a diversity of landscapes. The upshot of the dynamic and frequent LULCC process is fragmented land cover. This work's primary objective was to clarify how various machine learning algorithms performed in three different spatial and multispectral satellite image categorization tasks among specialists from both urban and rural locations. The most efficient algorithms for LULC classification have been found by performing atmospheric and geometric correction using a set of moderate and higher resolution images (Landsat-8, Sentinel-2, and Planet images) with similar phenological phases [17].

The lack of a single categorized high-resolution dataset with many class labels has also impeded the advancement of satellite image analytics. Two novel satellite datasets, SAT-4 and SAT-6, will be presented first. Secondly, a classification method that extracts topographies from a participation image, regularizes them, and provides the standardized feature directions to a deep belief network for cataloguing will be projected. Our system achieves a 97.95% categorization accuracy on the SAT-4 dataset, outperforming sophisticated object recognition techniques. The suggested approach achieved a 93.9% classification accuracy on the SAT-6. The superiority of unsupervised learning over traditional supervised learning techniques was demonstrated by related models using a random forest classifier. A numerical study based on the intrinsic dimensionality approximation and distribution separability principle validates the value of the suggested approach in acquiring better representations for satellite images [18].

The classification of satellite images using a perceptron neural network with the learning rule LEARNPN and the transfer function HARDLIM is shown in this paper. Earlier classification, the improved images are separated into a number of blocks and feature mining is agreed out by principal components analysis (PCA). As color plays a significant part in distinguishing the substances in the satellite images, color evidence is used in mining important topographies. Fifty images from Landsat are utilized for training and testing of the outcomes. The substances in the classes of water, land and vegetation are recognized based on red green blue (RGB) components. It is determined that selecting a suitable block size affects the categorizing correctness [19].

Convolutional neural networks have been successfully applied to multimedia techniques and used to create a scheme that can handle categorization without the need for human participation. Deep learning-based operational methods for classifying satellite images that leverage convolutional neural networks for feature extraction employing pretraining techniques from AlexNet, visual geometry group-19 (VGG-19), GoogLeNet, and ResNet50. Using three independent datasets (SAT4, SAT6, and UC Merced Land), the ResNet50 strategy achieves a more favorable result than other methods [20].

A PC-based incremental system is suggested in this study. For each image set, the edges are detected using the Canny operator on the panchromatic images, and to identify the vegetated areas to be used for masking, the normalized difference vegetation index (NDVI) is applied. Then, the discovered edges are improved using the edge thinning and division technique. Roads are automatically or semi-automatically derived from satellite images [21].

The ability of model stacking to lessen the effects of model bias and overfitting is one of main benefits. Final forecast is less inclined to be impacted by flaws or biases of a single model when numerous models projections are integrated. It can assist to capture a wider variety of information and increase overall prediction accuracy. Figure 1 represents the flow of a Blending approach where any four ML algorithms are considered as base models and one model is treated as a meta model. There are training and testing sets inside the data set. The identical training set and testing set are used to train and test each of the four base models. Once the test set is input into these base models, four different predictions P1, P2, P3 and P4 will be generated. These predictions will be merged and created as a meta training set on which the meta model (second level algorithm) will be trained. Finally, the test set is given as input to the meta model for making predictions. These predictions will be relatively accurate than the predictions done by individual base models.

There is a flaw in the blending approach. As the base models and meta model are trained on the same data set and because of overfitting and data leakage, the model may perform well on training sets but poorly on test sets. This flaw can be rectified using the holdout approach.

The dataset is first split into training and testing subsets in the holdout method. A training set and a validation set are further divisions of the training set. The various ML algorithms of the base model stack are trained using the training subset. The next step is to apply the validation set to the trained base models. These models will make some predictions. These predictions will be used as a training set for the meta model. The test set from the first step will be used for evaluating the blending model. The creation of the prediction set from the base models is shown in Figure 2. The blending strategy, shown in Figure 3, involves training a meta model on top of the underlying models' predictions. As an assessment set, the test set produced during the first data set split will be utilized. Blending uses a "one-holdout set", or a small portion of the training data (validation) to make predictions that will be "stacked" to form the training data of the meta-model, whereas stacking uses a k-fold validation scheme to generate the meta-model. In order to create the meta-model test data, predictions are also formed from the test data.

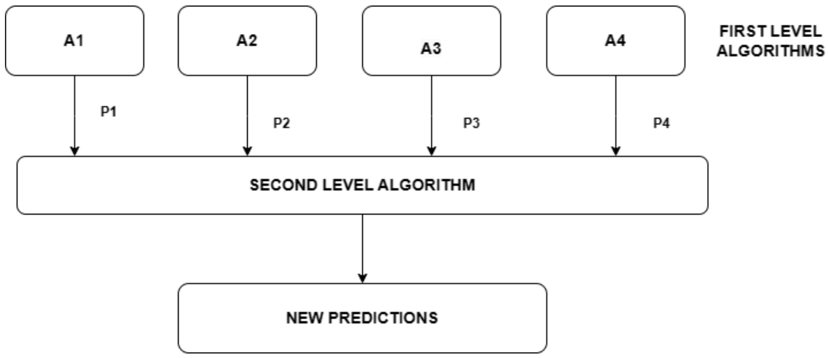

Figure 1. Blending approach

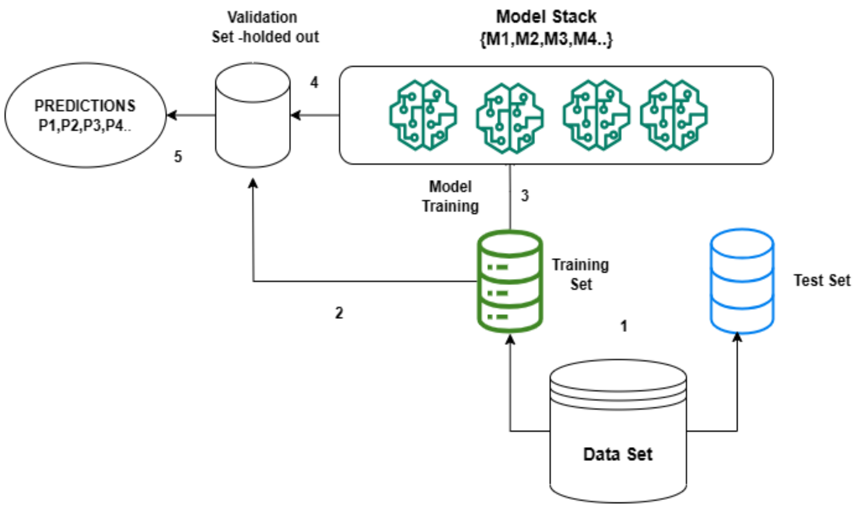

Figure 2. Generation of prediction from the base models

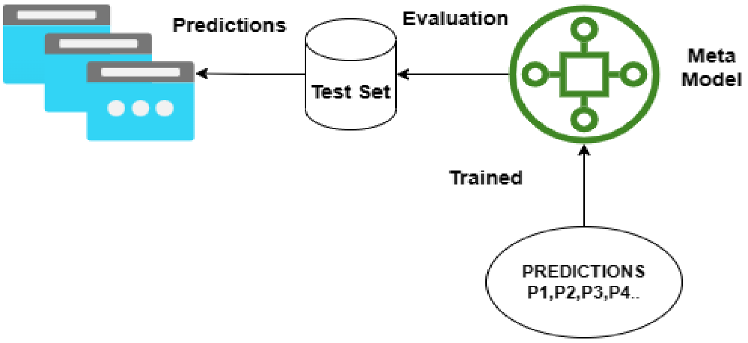

Figure 3. Blending using holdout approach

2. METHOD

2.1. Image classification using machine learning

For image categorization problems, machine learning methods have been frequently employed. Some of the common machine learning techniques used in image categorization include supervised learning, unsupervised learning, and deep learning. The most popular method for classifying images is supervised learning, in which a computer learning model is trained on a labelled dataset. By minimizing a loss function, the model in this method learns to map input pictures to their associated labels. DT, logistic regression (LR), SVM, and neural networks are a few of the well-liked supervised learning methods used in image categorization. Following algorithms were implemented in this paper:

a. Support vector machines (SVM): SVMs' capacity to generalize successfully even with few training data makes them very useful in the field of remote sensing. There are several kernel types in SVM:
   - We employed the radial basis function (RBF) kernel with parameter $gamma = 0.1$, $c = 0.01$.
   - We also implemented SVM using two other kernels: ‘Linear’ and ‘Polynomial’ [22].
b. K-nearest neighbors (KNN): KNN merely memorizes the training data and applies it to forecast new data points. This algorithm was implemented with a parameter ‘$n\_neighbors$’ of value 400.
c. Decision trees (DT): DT work by repeatedly dividing the data into the features that are most relevant to the job at hand, until a halting requirement is met. We used this algorithm with the parameters: $max\_depth = 4$, $max\_leaf\_nodes = 4$.

2.1.1. Image classification using ensemble learning

By merging many models or classifiers, ensemble learning is a potent technique for enhancing the accuracy and resilience of image classification systems. By mixing many models with various interpretability traits, ensemble learning may also be utilized to enhance the interpretability of image classification models. For instance, one model may be built to make predictions that are more accurate, while another model might be built to provide users more information about how a categorization choice was reached. The following classifiers were used:

a. Gradient boosting classifier: It is a powerful classifier created by combining many weak classifiers. The GBC technique builds a weak beginning classifier, like a decision tree, then repeatedly strengthens it by adding more weak classifiers that concentrate on the samples that the original classifiers misclassified. One of the GBC model's hyperparameter: ‘$n\_estimators$’, was given a value of 100 and other hyperparameters were also carefully adjusted to prevent overfitting.
b. Histogram-based gradient boosting classifier: Decision trees are iteratively added to the model in order for the histogram-based gradient boosting classifier to function.
c. Random forest classifier: To provide a final forecast, random forest constructs several decision trees and combines their predictions. To lessen overfitting and boost generalization, each tree is constructed using a random portion of the training data and a random subset of the features.
d. Bagging classifier: An ensemble classifier that employs bagging to boost performance is known as a bagging classifier. A decision tree is often the base learner in a bagging classifier, however other learner types may also be utilized. We have used k neighbor classifier as the base estimator for this ensemble algorithm with ‘$n\_estimators = 200$’. The forecasts of all the models are averaged to provide the final projection.
e. XGBoost classifier: The XGBoost method builds an ensemble of decision trees, each one created sequentially to fix the flaws of the one before it. This strategy assists in reducing the loss function, which enhances the model's general accuracy.

f. Stacking ensemble model: To determine how to combine the predictions from two or more underlying machine learning algorithms, stacking employs a meta-learning algorithm. For the base models that are required for stacking, we have combined the above mentioned five ensemble algorithms and used ‘$LogisticRegression$’ as the meta-learner model. In comparison to individual learners, the meta-model delivered even better accuracy by combining the predictions of the base models.

Most of classical machine learning algorithms suffer from scoring optimal classification performance over multi-spectral images. In this study, we propose stack-based ensemble-based learning approach to optimize classification performance [23]. Aiming at evaluating the advantages of classifier ensemble strategies and object-based image analysis (OBIA) method for satellite data classification under complex urban area [24]. It would be vitally necessary to implement machine learning methods for classification purposes to support the graphics processing unit (GPU) systems to work faster [25].

2.1.2. Dataset and experimental setup

The dataset selected for this research is satellite image classification dataset-RSI-CB256 [26]. This dataset has a total of 5,361 images divided into 4 classes mixed from Sensors and google map snapshot. The images are classified into cloudy (1,500), desert (1,131), green_area (1,500) and water (1,500). Figure 4 depicts sample images from the dataset. The data set was divided into training and testing sets in 80:20 ratio and for holdout it was divided into 70:15:15 for training, validation and testing.

In order to classify these images using machine learning and ensemble learning algorithms, we need a certain set of features for the provided data. Hence, we have applied first order statistical feature extraction by considering mean and standard deviation as main features for the provided data. Finding patterns and connections in the data that may be utilized to generate precise predictions or classifications is the aim of statistical feature extraction.

Figure 4. Sample images from data set includes water, green_area, cloudy and desert (left to right)

To develop the models, Google Colaboratory environment was used [27]. This platform is employed because it offers free high-end GPUs that can be used to process images swiftly and train models at a faster pace. After opening Google Colab in the browser, we can go to ‘Runtime’ → ‘Change runtime type’ → NVIDIA Tesla K80 GPU to enable the processor. The initial dataset is divided for 80% Training and 20% Testing. We have applied SVM classification algorithm, KNN classification algorithm and DT machine learning algorithms for specific parameters on the Satellite RSI dataset. Later, the ensemble learning algorithms, GBC, histogram-based GBC, random forest classifier, bagging classifier, XGBoost classifier were also applied to this dataset. All the ensemble learning algorithms were provided to the stacking ensemble learning model and integrated together to improve the overall accuracy that is provided by the separate ensemble algorithms. Blending using holdout approach was performed in three models by keeping the base model stack same and changing the Meta model.

a. Blender 1 (logistic regression as Meta model)

Gradient boosting, histogram-based gradient boosting classifier, random forest, bagging classifier and XGBoost classifier were considered as the base models. Logistic regression was considered as meta model. These models were hyperparameter tuned using *GridSearchCV*.

Step 1: Split the data into training and test sets
Step 2: Split the training data into training and validation sets
Step 3: Train the base models with hyperparameter tuning using *GridSearchCV* on the training set
Step 4: Make predictions with the base models on the validation set
Step 5: Combine predictions from the base models to create a new feature matrix for meta-model
Step 6: Train the meta-model (LR) on the new feature matrix with hyperparameter tuning.
Step 7: Make predictions with the blending model on the test set

Step 8: Combine the predictions from the base models on the test set to create a new feature matrix for the meta-model
Step 9: Make predictions with the blending model (using base models' predictions as input to the meta-model) on the test set
Step 10: Evaluate the blending model's performance on the test set

b. Blender 2 (Random forest as Meta model)

Gradient boosting, histogram-based gradient boosting classifier, random forest, bagging classifier and XGBoost classifier were considered as the base models. Random forest was considered as Meta model. These models were hyperparameter tuned using *GridSearchCV*.

Step 1: Split the data into training and test sets.
Step 2: Split the training data into training and validation sets.
Step 3: Train the base models with hyperparameter tuning using *GridSearchCV* on the training set.
Step 4: Make predictions with the base models on the validation set.
Step 5: Combine predictions from the base models to create a new feature matrix for meta-model.
Step 6: Train the meta-model (random forest) on the new feature matrix with hyperparameter tuning.
Step 7: Make predictions with the blending model on the test set.
Step 8: Combine the predictions from the base models on the test set to create a new feature matrix for the meta-model.
Step 9: Make predictions with the blending model (using base models' predictions as input to the meta-model) on the test set.
Step 10: Evaluate the blending model's performance on the test set.

c. Blender 3 (XGBoost as the Meta model)

Gradient Boosting, histogram-based gradient boosting classifier, random forest, bagging classifier and XGBoost classifier were considered as the base models. XGBoost was considered as Meta model

Step 1: Split the data into training and test sets.
Step 2: Split the training data into training and validation sets.
Step 3: Train the base models with hyperparameter tuning using *GridSearchCV* on the training set.
Step 4: Make predictions with the base models on the validation set.
Step 5: Combine predictions from the base models to create a new feature matrix for meta-model.
Step 6: Train the meta-model (XGBoost) on the new feature matrix with hyperparameter tuning.
Step 7: Make predictions with the blending model on the test set.
Step 8: Combine the predictions from the base models on the test set to create a new feature matrix for the meta-model.
Step 9: Make predictions with the blending model (using base models' predictions as input to the meta-model) on the test set.
Step 10: Evaluate the blending model's performance on the test set.

3. RESULTS AND DISCUSSION

The paper performed satellite image classification by creating solo models as well as ensemble models. When evaluated on the 5,361 images of the RSI-CB256 dataset, SVM (RBF) model achieved 72.84%, SVM (linear) model achieved 73.38%, SVM model 68.32 %, KNN model 80.56% and DT model achieved 75.33 % respectively. Table 1 represents the accuracy of solo models created from SVM (RBF), SVM (linear), SVM (Polynomial), KNN and DT. In the solo models KNN based model outperforms the other models with an accuracy of 80.56%. Figure 5 visualizes Table 1. SVM (polynomial) is the worst performer with 68.32% succeeded by SVM (RBF) 72.84%, SVM (linear) 73.38%, DT 75.33% succeeded by KNN 80.56% which is the best performer of the lot.

Table 1. Results obtained from solo machine learning algorithms

S. No	Algorithm	Accuracy (In %)
1	SVM (RBF)	72.84
2	SVM (linear)	73.38
3	SVM (Polynomial)	68.32
4	KNN	80.56
5	Decision Tree	75.33

In the second set of experiments this paper applied the ensemble techniques on the dataset and achieved results as shown in Table 2. A total of nine ensemble models were designed for classifying satellite images. The gradient boosting classifier achieved an accuracy of 86.33% in classifying the satellite images.

Histogram variant of the gradient boosting classifier classified the images with an accuracy of 87.84%. The random forest-based model identified 84.20% of the images that were input to it correctly. The bagging ensemble model recognized 88.90% of the images accurately. The XGBoost model achieved an accuracy of 87.84% accuracy whereas the stack model achieved 88.19% accuracy. Clearly the blended models outperform other normal ensemble models including the stack model by showing an accuracy rate of 93.7%, 94.14% and 93.87% respectively. Figure 6 is a graphical representation of the rate of accuracy achieved from each model. Out of the three blended models, blender 2 that uses random forest as Meta model stands top with an accuracy rate of 94.14 followed by blender 3 (XGBoost Meta model) and blender 1 (logistic regression Meta model). All experiments were performed using Google Colaboratory environment using python.

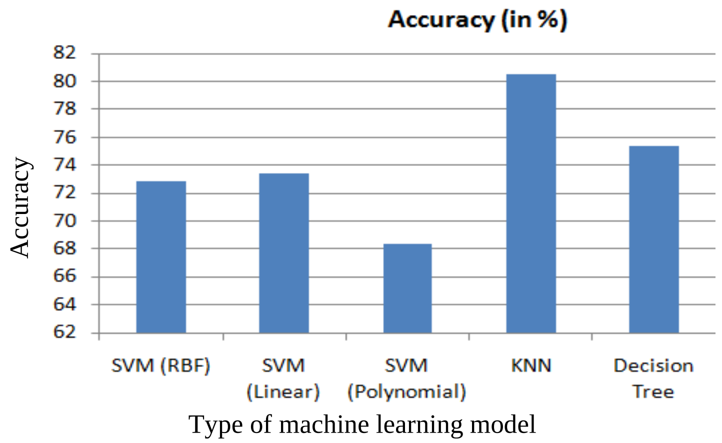

Figure 5. Performance of machine learning models

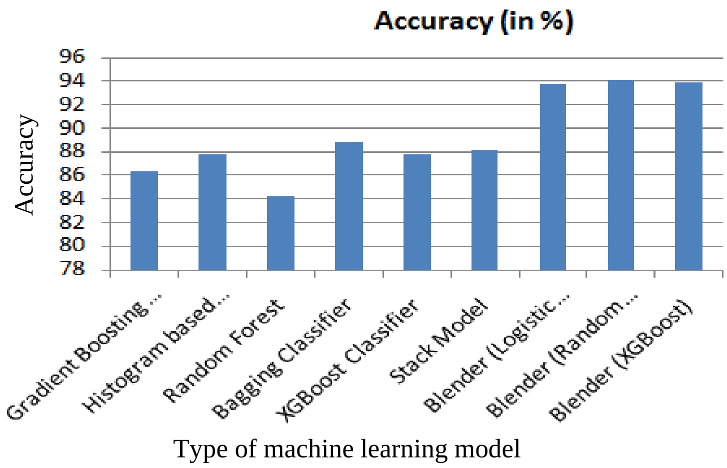

Figure 6. Accuracy of ensemble learning algorithms

Table 2. Results obtained from ensemble learning algorithms

S. No	Model	Accuracy (in %)
1	Gradient boosting classifier	86.33
2	Histogram-based gradient boosting classifier	87.84
3	Random forest	84.20
4	Bagging classifier	88.90
5	XGBoost classifier	87.84
6	Stack model	88.19
7	Blender 1	93.70
8	Blender 2	94.14
9	Blender 3	93.87

4. CONCLUSION

Now-a-days accurate classification of satellite images is a critical task as weather forecasting, mining, military intelligence, real estate and routing are dependent on this area. With the increase in computing power machine learning algorithms are proving to be a big boon in this field. This paper created two sets of satellite image classification models namely solo and ensemble models. The experiments conducted by this research revealed that even though the solo models require less time to classify, they lack accuracy when compared to ensemble models. In the ensemble models blending models with handout approach perform well when compared with the boosting and stack models. The blender model 2 that uses Random Forest as meta model outperformed all other models with an accuracy of 94.14%. In future this work plans to use different datasets to prove the efficiency of blending methods. Even deep learning methods can be used to classify the satellite images.

ACKNOWLEDGEMENTS

The authors express their gratitude and acknowledge the support provided by Dayananda Sagar College of Engineering.

REFERENCES

[1] N. Abdul Azeem, S. Sharma, and S. Hasija, "Classification of satellite images using an ensembling approach based on deep learning," *Arabian Journal for Science and Engineering*, vol. 49, no. 3, pp. 3703–3718, Mar. 2024, doi: 10.1007/s13369-023-08143-7.

[2] D. Guo, Z. Wu, Y. Zhang, and Z. Shen, "Semi-supervised remote sensing image scene classification based on generative adversarial networks," *International Journal of Computational Intelligence Systems*, vol. 15, no. 1, Oct. 2022, doi: 10.1007/s44196-022-00150-0.

[3] S. Puttinaovarat, K. Khaimook, and P. Horkaew, "Land use and land cover classification from satellite images based on ensemble machine learning and crowdsourcing data verification," *International Journal of Cartography*, pp. 1–21, Mar. 2023, doi: 10.1080/23729333.2023.2166252.

[4] K. Thiagarajan, M. Manapakkam Anandan, A. Stateczny, P. B. Divakarachari, and H. K. Lingappa, "Satellite image classification using a hierarchical ensemble learning and correlation coefficient-based gravitational search algorithm," *Remote Sensing*, vol. 13, no. 21, Oct. 2021, doi: 10.3390/rs13214351.

[5] M. J. Horry, S. Chakraborty, B. Pradhan, N. Shulka, and M. Almazroui, "Two-speed deep-learning ensemble for classification of incremental land-cover satellite image patches," *Earth Systems and Environment*, vol. 7, no. 2, pp. 525–540, Jun. 2023, doi: 10.1007/s41748-023-00343-3.

[6] S. Borra, R. Thanki, and N. Dey, "Satellite image classification," in *Satellite Image Analysis: Clustering and Classification*, 2019, pp. 53–81.

[7] R. Sonobe, Y. Yamaya, H. Tani, X. Wang, N. Kobayashi, and K. Mochizuki, "Crop classification from sentinel-2-derived vegetation indices using ensemble learning," *Journal of Applied Remote Sensing*, vol. 12, no. 02, May 2018, doi: 10.1117/1.JRS.12.026019.

[8] N. Kobayashi, H. Tani, X. Wang, and R. Sonobe, "Crop classification using spectral indices derived from Sentinel-2A imagery," *Journal of Information and Telecommunication*, vol. 4, no. 1, pp. 67–90, Jan. 2020, doi: 10.1080/24751839.2019.1694765.

[9] T. S. Asawa, V. Balaji, and T. Helwatkar, "Deep ensemble learning for agricultural land mapping and classification from satellite images," *International Journal of Engineering Research & Technology (IJERT)*, vol. 10, no. 5, pp. 1063–1069, 2021.

[10] B. Ekim and E. Sertel, "Deep neural network ensembles for remote sensing land cover and land use classification," *International Journal of Digital Earth*, vol. 14, no. 12, pp. 1868–1881, Dec. 2021, doi: 10.1080/17538947.2021.1980125.

[11] Y. Zhang, J. Liu, and W. Shen, "A review of ensemble learning algorithms used in remote sensing applications," *Applied Sciences*, vol. 12, no. 17, Aug. 2022, doi: 10.3390/app12178654.

[12] H. Ebrahimy and Z. Zhang, "Per-pixel accuracy as a weighting criterion for combining ensemble of extreme learning machine classifiers for satellite image classification," *International Journal of Applied Earth Observation and Geoinformation*, vol. 122, Aug. 2023, doi: 10.1016/j.jag.2023.103390.

[13] B. N. Pandey, A. K. Shrivastava, and A. Rana, "Deep learning based ensemble model for satellite image classification," *TechRxiv*, 2023.

[14] A. K. Rai, N. Mandal, A. Singh, and K. K. Singh, "Landsat 8 OLI satellite image classification using convolutional neural network," *Procedia Computer Science*, vol. 167, pp. 987–993, 2020, doi: 10.1016/j.procs.2020.03.398.

[15] N. Manohar, M. A. Pranav, S. Aksha, and T. K. Mytravarun, "Classification of satellite images," in *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2020*, 2021, pp. 703–713, doi: 10.1007/978-981-15-7078-0_70.

[16] P. Iyer, S. A, and S. Lal, "Deep learning ensemble method for classification of satellite hyperspectral images," *Remote Sensing Applications: Society and Environment*, vol. 23, Aug. 2021, doi: 10.1016/j.rsase.2021.100580.

[17] A. Rahman *et al.*, "Performance of different machine learning algorithms on satellite image classification in rural and urban setup," *Remote Sensing Applications: Society and Environment*, vol. 20, Nov. 2020, doi: 10.1016/j.rsase.2020.100410.

[18] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, "DeepSat: a learning framework for satellite imagery," in *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Nov. 2015, pp. 1–10, doi: 10.1145/2820783.2820816.

[19] M. S. Devi and S. Chib, "Classification of satellite images using perceptron neural network," *International Journal of Computational Intelligence Research*, vol. 15, no. 1, pp. 1–10, 2019.

[20] M. A. Kadhim and M. H. Abed, "Convolutional neural network for satellite image classification," in *Springer Asian Conference on Intelligent Information and Database Systems*, 2020, pp. 165–178, doi: 10.1007/978-3-030-14132-5_13.

[21] S. K. L. Yongguan Xiao, "Feature extraction using very high resolution satellite imagery," in *IEEE International IEEE International IEEE International Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004*, vol. 3, pp. 2004–2007, doi: 10.1109/IGARSS.2004.1370741.

[22] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: a review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259, May 2011, doi: 10.1016/j.isprsjprs.2010.11.001.
[23] T. Aboneh, A. Rorissa, and R. Srinivasagan, "Stacking-based ensemble learning method for multi-spectral image classification," *Technologies*, vol. 10, no. 1, Jan. 2022, doi: 10.3390/technologies10010017.
[24] R. Han, P. Liu, G. Wang, H. Zhang, and X. Wu, "Advantage of combining OBIA and classifier ensemble method for very high-resolution satellite imagery classification," *Journal of Sensors*, vol. 2020, pp. 1–15, Nov. 2020, doi: 10.1155/2020/8855509.
[25] H. Ferdous, T. Siraj, S. J. Setu, M. M. Anwar, and M. A. Rahman, "Machine learning approach towards satellite image classification," in *Proceedings of International Conference on Trends in Computational and Cognitive Engineering. Advances in Intelligent Systems and Computing*, 2021, pp. 627–637.
[26] M. Reda, "Satellite image classification," Kaggle, 2022. https://www.kaggle.com/datasets/mahmoudreda55/satellite-image-classification (accessed Feb. 3, 2022).
[27] V. Lall, "Google Colab – the beginner's guide," *Medium*, 2018. https://medium.com/lean-in-women-in-tech-india/google-colab-the-beginners-guide-5ad3b417dfa (accessed Apr. 05, 2022).

BIOGRAPHIES OF AUTHORS

Suresh Kumar Musali holds a M.Tech degree in computer science and engineering from Visvesvaraya Technological University, Belagavi. He currently serves as an assistant professor in Dayananda Sagar College of Engineering, Bangalore. He has 16 years of experience in teaching. He has published around 10 papers in journals and conferences. He can be contacted at email: sureshkumar-ise@dayanandasagar.edu.

Rajeshwari Janthakal working as professor and Head of Information Science and Engineering Department, Dayananda Sagar College of Engineering, Bangalore. She earned her Ph. D degree from Visvesvaraya Technological University, Belagavi. She has 21 years of experience. She has published around 30 papers in journals and conferences. She has also published two patents and received a Research grant from Indian Council of Medical Research. She can be contacted at email: rajeshwarij-ise@dayanandasagar.edu.

Nuvvusetty Rajasekhar currently working as professor in the Department of Information Technology, Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad. He received a doctoral degree in the year 2015 from Acharya Nagarjuna University, Guntur and M.Tech in Computer Science and Engineering (2006) from Bharath University, Chennai. He graduated in mechanical engineering from S.V. University, Tirupati. He is currently, India. He has a total of 18 plus years' experience in teaching. He is actively involved in research and some of his research interests include network security, data mining, machine learning and artificial intelligence. He has published more than 30 research papers in international journals and conferences and has received best paper award in international conference at IEEE Conference at Agadir, Morocco in the year 2016. He is also serving as reviewer and editorial board member for various reputed peer reviewed journals and technical committee member for various international conferences. He can be contacted at email: rajasekhar531@gmail.com.