International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Classification Performance of Land Use from Multispectral Remote Sensing Images using Decision Tree, K-Nearest Neighbor, Random Forest and Support Vector Machine Using EuroSAT Data

Reena Thakur[1] and Prashant Panse[2]

Submitted: 06/06/2022 **Accepted:** 10/09/2022

Abstract— Remote sensing is commonly used in remote sensing applications for land cover and land use classification using remotely sensed data. Different algorithms for LULC mapping need to be compared to determine which one is most accurate for further use of Earth observations. In this study, performance of four machine learning algorithms, specifically decision tree (DT), k-nearest neighbor (KNN), support vector machine (SVM), and random forest (RF) was examined with the help of satellite images from the EuroSAT dataset. Accuracy assessment was performed using the training, testing, and validation methods. With the help of the confusion matrix, the classification output, the prediction test, and validation accuracy assessment were assessed to obtain the classifier with more accuracy. Validating the classification findings against actual data would reveal the optimal performance. According to the EuroSAT dataset, the overall classification accuracy was 98.57 percent, which is higher than the K-nearest neighbor classifier and more suitable for satellite image classification. Appropriate LULC maps can be produced by accurately classifying the data. This map can be used in variety of applications. Based on Sentinel-2 satellite photos, we provide a new dataset with 27,000 classified images from 13 spectral bands and 10 classifications. The suggested research's categorization approach opens the door to an extensive range of Earth observation applications. Maps may be improved by using a categorized system, which we illustrate here.

Keyword: *machine learning algorithm; land use/land cover (LULC); decision tree; k-nearest neighbor; support vector machine; random forest.*

I Introduction

Observing and documenting the Earth's surface via the use of a variety of technologies is the process of remote sensing (RS). As a primary source of remotely sensed data, satellite images are now widely used for a variety of purposes, including land use change monitoring, forestry, and more. The purpose of remote sensing is to classify land surface characteristics by using satellite images [2] for land use and land cover planning.

RS applications rely heavily on image analysis and pattern recognition [3]. It is possible to evaluate the classification itself in some circumstances. As an example, a map like this one for land use categorization may be produced using remotely sensed data. Digital image classification relies heavily on image categorization as a technique. Images are classified using algorithms implemented in software, known as classifiers. Various image classification methods are now being employed by researchers for a variety of reasons [4].

There has been a convergence in coming years between the remote sensing and data science fields due to several issues. There are several well-known data science competitions that use advanced machine learning techniques, such as Kaggle (Google 2019). In Kaggle, companies provide their data for contenders to use as training input for their models, and the winners receive monetary rewards. The general public has been provided free access to the data of several new Earth-observing satellites in recent years. The most recent has been EuroSAT Data.

For this research, we will utilize the EuroSAT dataset and four different classification approaches (decision trees (DTs), k-nearest neighbors (kNN), support vector machine (SVM), and random forest (RF), to assess classification strategies for identifying land use and land cover (KNN).

We classified 27,000 geo-referencing photos of land cover and land use from the EuroSAT dataset. There is a public repository for the EuroSAT data at

[1] *en19cs601002@medicaps.ac.in, prashant.panse@medicaps.ac.in*
orcid.org/0000-0002-7595-6934, orcid.org/0000-0001-8129-1214
Department of Computer Science and Engineering
Research Scholar[1], Medi-Caps University, Indore[1,2]

https://github.com/phelber/EuroSAT [4] [5]. Extensions are made to decision trees, k-nearest neighbors, support vector machines, random forests (RF), to increase classification accuracy. The categorization results of all the approaches were compared in this study.

As far as machine learning algorithms go, random forest is one of the most extensively used. It may be employed with categorical or continuous variables, which is why classification and regression analysis rely on this approach so heavily [16]. Land Use [17]-Land Cover [18] applications have benefited from RF's versatility. RF was compared to the DTs and determined to be the most accurate, thereby exceeding the DTs.

The Support Vector Machine (SVM) is considered the most accurate classifier because of its overall accuracy. Remotely sensed data categorization can benefit greatly from the use of SVM. When compared to DTs and k-NN, this method has superior overall accuracy. Training imbalance may have an impact on this categorization accuracy [20].

In this section of the paper, everything is arranged as follows: Section II provides a brief overview of other works, their context, and their goals. Machine learning algorithms are discussed in Section III. Section IV goes into great depth about the suggested approach. Section V is all about the outcomes and evaluations. Result and performance analysis is described in Section VI.

II General Approach of the research

No prior information about the area classes is considered in traditional classification techniques; only information pertaining to training and testing venues is. It's impossible to test the idea since, for example, there is no data on metropolitan regions. This study examines whether or not incorporating past land-cover information to classify a picture may enhance classification accuracy. It is then weighted according to the section into which the class falls. Map data defines a built-up land-cover class for those pixels that fall outside of an already established built-up region.

Background

A. Use and Cover of the Land among the many different types of land cover, there is surface cover, which includes annual crops and forests, as well as herbaceous plants and highways, as well as subsurface features such as lakes and rivers, as well as subterranean features such as grazing and permanent crops. Forests, grasslands, croplands, wetlands, and urban structures are all types of land cover. However, human land use refers to the control and alteration of the usual environment by humans, such as urban development. As a result, both human-made and natural land cover are depicted in the LULC diagram. In some cases, there is no clear connection between land cover and land use. It is possible to maintain multiple types of land cover on one piece of land [6].

B. Classification of an image is a computer-aided approach for gathering data for digital image processing. There are typically four phases involved in image classification:

1. Preprocessing: reducing the haze, identifying the band ratios, and correcting the atmosphere.
2. It's the process of selecting a specific feature to characterize the pattern that is called a training sample.
3. Testing is the third step in figuring out the best way to compare the target to the picture pattern.
4. Validation and evaluation of the categorical accuracy of the picture [7].

Using the accuracy evaluation, the classification results are subject to an accuracy assessment. It is used to make the confusion matrix or error matrix, which is used to check the accuracy of classification results for all classifiers.

Objective

1. Researchers interested in land use/land cover categorization can utilize this paper's goal as a road map to help them anticipate which techniques and geographic regions they should focus on next.
2. A full comparison of detailed methodologies and advantages is presented in this work, which begins with picture preprocessing, acquisition, and ends with analysis and validation of the classification process.
3. The LU/LC image categorization process is explained in detail in this work.
4. There is a EuroSAT dataset used in this paper.

III Machine learning algorithms

This section presents a detailed account of the algorithms used in image classification. With the help of EuroSAT Dataset, the overall accuracy, the producer accuracy, and the user accuracy have been determined for all methods/algorithms used.

Support vector machines

Cortes and Vapnik (1995) has explained SVM algorithm at first [8] based on the survey by Vapnik (1982) [9]. In the field of Remote sensing applications generally uses supervised learning technique. The working of SVM algorithm is on the hyperplane (decision boundary) which helps to find the optimum minimization. In this way, the classification problem can be separated into classes that can be reliably classified using the training data.

As part of the algorithm, a hyperplane boundary is found iteratively in an n-dimensional classification space in order to differentiate patterns in the training data, then the configurability is applied to a distinct evaluation dataset.

Mountrakis, Jungho, and Ogole (2011) define dimensions simply as the number of vectors and spectral bands in a multiband composite [10].

Random forests

RF is based on the concept that a combination of bootstrap aggregated classifiers performs better than a DTs, supervised machine learning algorithm [11]. The bootstrap component, wherein each individual tree is parameterized with the help of an arbitrarily sampled set of notes and with replacement from the training data [12]. Multicollinearity reduced by de-correlating the trees. The proportion of observations which are not considered for this work are comprised in the assessment and are known as "out-of-bag" samples. A few of these decision trees models are built using alternative classifications of the input variables, with the approximate majority vote of each class summed through all trees as the final output.

Decision Trees

The Decision Tree classification method that determines the structure of a classification tree is a supervised classification technique. Top or root nodes in DT are responsible for including the functions and attributes required to perform the categorization test. The tree edged that represents the test result, and the leaf nodes, hold the mark for the class. Han J. and Kamber (2006) in their book presented that no knowledge base or additional parameters are needed for DT generation and can be efficiently represented and understood by a human. In addition, the authors explained that the DT learning and classification method is easy to use and effective in helping achieve good accuracy in classifications. Classification decisions may be based on decision rules or on binary tree structures. In order to carry out the SAR classification, knowledge-based or rule-based methods were used previous by the author (Pierce et al. 1997). In order to discriminate against major classification regions such as urban, high vegetation, short vegetation, and bare soil, texture details and the backscatter coefficient values of both L and C-band polarizations were used.

A decision tree is a classification system that uses specifically a set of rules at each node in the tree to iteratively partition a dataset into subcategories. It has gained popularity in the classification of land cover with the use of remote sensing data because it offers various advantages over traditional supervised classification processes such as maximum likelihood classification. A decision tree can be used to describe decision and decision making effectively and clearly in decision analysis. Basically, a decision tree is a tree-shaped graphic that is used to make such decisions or to display statistical probability. The tree is designed to demonstrate how and why one decision may leads to the next, with the branches indicating that each alternative is mutually incompatible [15]. The decision tree's weakness is high overfitting in training data, which means that a small change in data can result in a completely different tree.

K-nearest neighbor

As a supervised algorithm, kNN is mostly used for classification and regression [13]. Training samples considered as input are the k-nearest neighbors' training samples, and the output will be assumed to be a class, which will be classified based on the majority vote of its neighbors, the most frequent class among the k nearest neighbors assigned as the Classifier. The basic idea is of selecting the nearest training sample and then determine its level depending on the distance. The significant drawback of this technique is that it is difficult to estimate the value of parameter k. Furthermore, the technique does not specify which sort of distance to utilize or which characteristic to employ to achieve the best outcomes. It has been extensively studied how decision trees perform in machine learning and data mining. They are nonparametric supervised learning methods that make it easier to perform classification and regression [14].

IV Methods

A. Research structure

As shown in Fig. 1, the methodology structure in this study is divided into three steps. Firstly, EuroSAT dataset is to be taken as input and preprocessing of the satellite images is required to remove the noise and image enhancement. Secondly, the dataset is divided into three main samples, train, test, and validate samples on which the SVM, DTs, RF and k-NN are applied. Before this the training model is to be built once to reduce the training process again and again. Whereas in third step Land Use/ Land Cover classification is to be done followed by the accuracy assessment for each classifier. The evaluation of image classification has been calculated from confusion matrix with the help of producer's accuracy, user's accuracy, overall accuracy.

B. Dataset Creation

Each day, over 1.6 terabytes of compressed images is provided by the Sentinel-2 satellite constellation. However, the lack of labelled ground truth data with the amount of data makes supervised machine learning less effective. EuroSAT dataset is a collection of images which comprised of 10 different classes, each comprising between 2000 and 3000 images. 27000 images make up this collection. They measure 64 pixels by 64 pixels. The ten categories of land use and land cover used in the analysis allowed us to generate thousands of image patches. Annual Crop, Permanent Crop, and Pastures are included in the proposed dataset to distinguish between diverse agricultural land usages. The dataset also discriminates against built-up regions, such as roadways, manufacturing facilities, and residential properties, since it covers groups. There are several bodies of water to choose from, including Lakes, Rivers, and the Sea. Forests and herbaceous vegetation, for example, are new unexplored ecosystems that each have their own class. Fig. 2 displays a list of classes impacted and four examples from each class. The EuroSAT land-use and land-cover classes are described below in table 1.

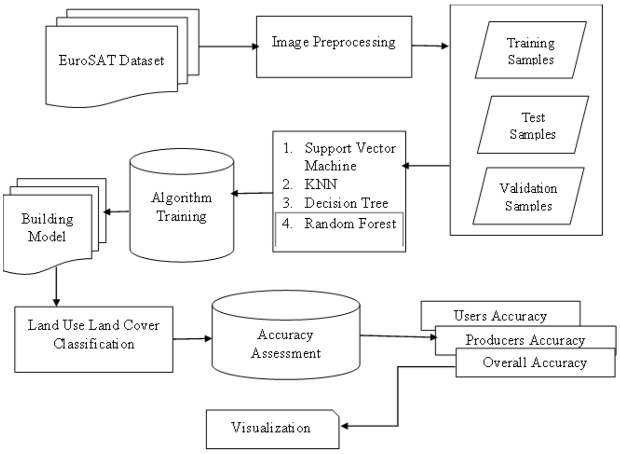


Figure 1- the methodology structure

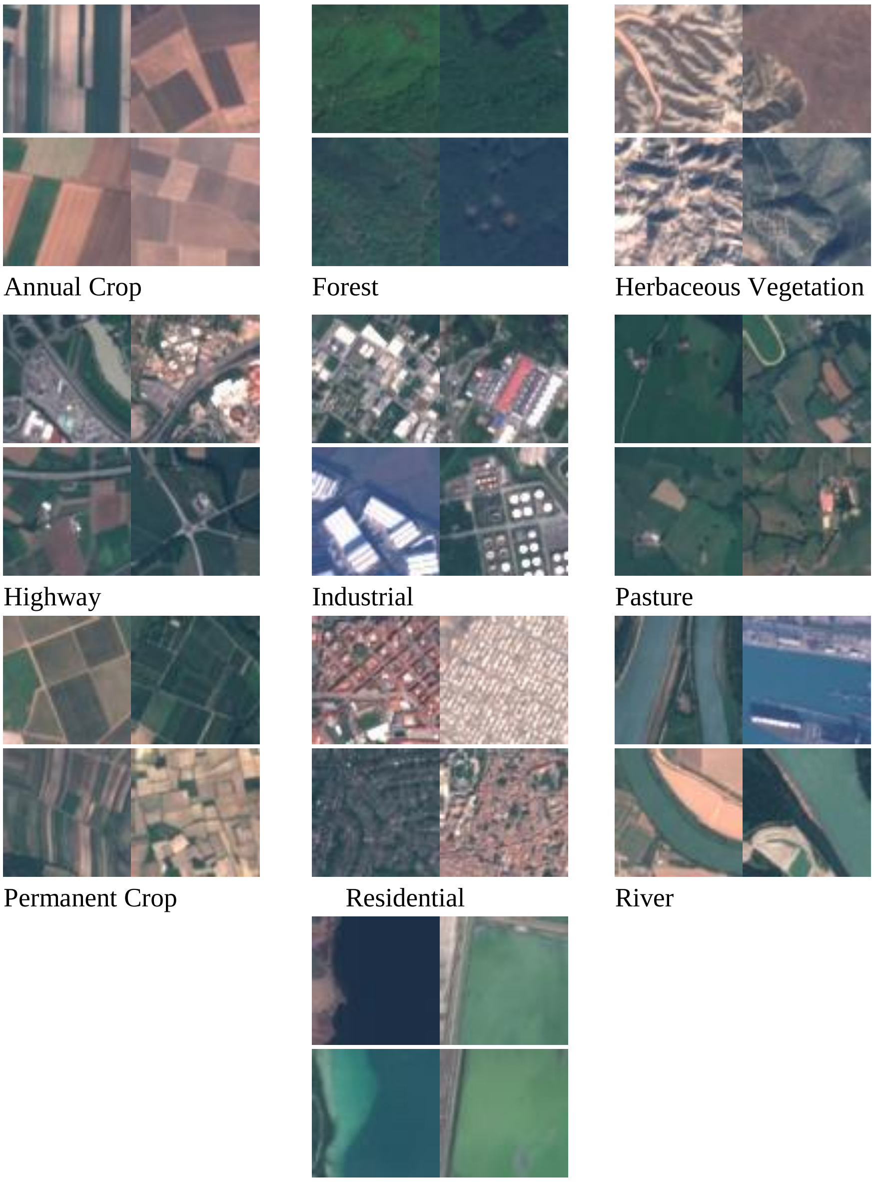
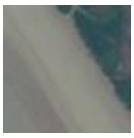


Figure 2- Sample image of EuroSat dataset [4].

Table 1 The EuroSAT land-use and land-cover classes are described.

	Classes	**Description of Class**	**Sample of Class**
	Permanent- Crop	The soil is permanently occupied by a crop, which produces harvests for several years in a row. Fruit orchards, vineyards, and olive groves, to name a few examples. Crops are harvested each year.	
	Annual-crop	Annual crops take up soil space and provide limited harvests. months. Barley, maize, etc., are examples of such crops.	
	Pastures	Domesticated livestock such as cattle, horses, sheep, goats, etc can be found in the pastures.	
	River	A river is a natural stream of water that travels toward its destination such as a lake, a sea, and an ocean	
EuroSAT Dataset	Sea &lake	An area that is completely encircled by water.	
	Forest	A large piece of land that has been heavily forested.	
	Herbaceous vegetation	Herbaceous plants, also known as vascular plants, are found in these areas. Plants with no above-ground woody stems.	
	Industrial building	A human-made environment used to protect industrial equipment.	
	Residential building	Human covered areas for their own sheltering.	
	Highway	Human covered areas for their travelling.	

C. Image Classification

Typically, image classification consists of four steps: first, the image is pre-processed, which includes removing noise, finding band ratios, and making atmospheric adjustments. Training sample refers to the selection of specific criteria features with which to describe the pattern based on the selection of the training sample. Additionally, it is important to choose the best method for comparing the target to the image pattern. Lastly, is the accuracy assessment of the image classification [22]. In this paper, supervised classification classifiers such as SVM, DTs, RF and k-NN have been used for land use and land cover classification to classify the image of the specific classes.

D. Validation Technique

To validate model performance four validation techniques overall accuracy, producer accuracy, and user accuracy [23].

The formula for overall accuracy is Overall accuracy = o / n. '

n = total number of test samples, o = number of accurately predicted samples observed.

Algorithm

```
   import accuracy_score, confusion_matrix

   #import matplotlib.pyplot as plt

   from imagesData import load_data


   while 1:
       try:
           with open('data..', 'rb') as f:
               print("\nLoading Saved data from file ...")
               a. images_train
               b. images_test
               c. images_val

               aa. targets_train
               bb. targets_test
               cc. targets_val

               load Labels
           break
       except:
           load_data(..)


def train():


    clf1 = tree.DecisionTreeClassifier()
    clf2 = neighbors.KNeighborsClassifier()
    clf3 = forest.RandomForestClassifier()
    clf4 = supposrt.SupportVectorMachineClassifier()

    clfs = [clf1, clf2, clf3, clf4]
    for clf in clfs:
        Training Model on {len_train_data} samples with {clf} ..
        clf.fit(x_train, y_train)


def checkForTrainedModels():
    global clfs

    try:
        Checking for Previously Trained Model ..
        with open("trained_model", "rb") as f:
            clf1 = pk.load(f)
            clf2 = pk.load(f)

        clfs = [clf1, clf2, clf3, clf4]
```

```
def test():
    global clfs, l

    x_test, y_test = images_test, targets_test
    acc = {}
    l = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

    len_test_data = len(y_test)
    for clf in clfs:
        Predicting test
        a = accuracy_score(y_test, prd)*100
        Prediction complete
        acc[clf] = a
        cm = confusion_matrix(y_test, prd)
        Display confusion matrix
        Get accuracy

def validate():
    global clfs, l

    print("-"*160)
    acc = {}
    len_val_data = len(targets_val)
    for clf in clfs:
        Validating
        prd = clf.predict(images_val)
        a = accuracy_score(targets_val, prd)*100
        acc[clf] = a
        #print(list(clf.predict_proba(images_val[:25])))
        cm = confusion_matrix(targets_val, prd)
        Display Confusion matrix
        Display validate accuracy
checkForTrainedModels()
test()
validate()
```

V Analysis and Result

In section V, the performance evaluation of the proposed model uses EuroSAT dataset for the classification of 10 LULC classes. These are highway, annual crop, herbaceous vegetation, forest, industrial building, pastures, river, permanent crop, residential building and sea and lake. This was accomplished by using the EuroSAT dataset, which includes 27,000 image patches, 70% of which were used to train, 30% for testing. Table 2 shows the results.

Table 2 Confusion Matrix- Test data results of classification algorithms where PA-Producer's Accuracy, UA-User's Accuracy

Algorithm	Decision Tree			k-NN			Support Vector Machine			Random Forest		
Class/Accuracy	PA	Producer's Total	UA	PA	Producer's Total	UA	PA	Producer's Total	UA	PA	Producer's Total	UA
Annual Crop	36.5	304	100%	26.47	661	100%	51.42	280	100%	68.92	379	100%
Forest	70.7	277	100%	0.0	20	100%	63.24	419	100%	65.15	348	100%
Herbaceous Vegetation	16.18	309	100%	10.00	230	100%	28.35	261	100%	26.05	213	100%
Highway	18.29	246	100%	68.75	16	100%	43.85	114	100%	55.26	140	100%
Industrial	49.46	188	100%	0	0	100%	74.49	298	100%	67.78	291	100%
Pasture	23.58	212	100%	6.60	106	100%	36.33	333	100%	33.03	232	100%
Permanent Crop	19.18	245	100%	0.0	3	100%	37.39	246	100%	13.04	169	100%
Residential	33.56	292	100%	0.0	0	100%	51.01	345	100%	48.98	386	100%
River	28.35	261	100%	70.00	50	100%	45.72	234	100%	56.83	262	100%
Sea Lake	65.02	366	100%	18.15	1614	100%	46.12	271	100%	94.09	274	100%

	a	b	c	d	e	f	g	h	i	j
a	111	2	49	26	16	15	37	20	16	8
b	2	196	3	5	0	16	1	1	6	70
c	37	5	50	38	19	40	37	35	27	12
d	29	6	31	45	14	24	17	29	50	5
e	25	0	21	25	93	4	33	36	11	2
f	12	16	39	15	0	50	10	19	23	16
g	37	0	55	19	18	10	47	42	19	3
h	23	4	33	33	21	15	42	98	28	3
i	24	17	26	38	7	28	16	11	74	9
j	4	31	2	2	0	10	5	1	7	238

Prediction on Test Data-Decision Tree

	a	b	c	d	e	f	g	h	i	j
a	175	1	34	1	0	12	0	0	4	73
b	0	0	0	0	0	0	0	0	0	300
c	63	3	23	0	0	13	0	0	1	197
d	67	4	48	11	0	21	1	0	9	89
e	176	0	25	2	0	8	2	0	0	37
f	4	0	16	0	0	7	0	0	0	173
g	98	0	22	0	0	5	0	0	1	124
h	41	0	21	1	0	9	0	0	0	228
i	35	10	40	1	0	29	0	0	35	100
j	2	2	1	0	0	2	0	0	0	293

Prediction on Test Data-k-NN

	a	b	c	d	e	f	g	h	i	j
a	193	0	22	10	14	12	11	13	22	3
b	0	273	2	0	0	19	0	0	2	4
c	47	7	68	14	13	36	27	50	32	6
d	27	4	12	63	15	16	17	47	49	0
e	9	0	0	6	202	0	7	25	1	0
f	0	18	41	6	0	110	3	13	5	4
g	66	0	30	10	14	9	60	56	5	0
h	13	0	16	16	29	3	43	169	11	0
i	18	20	18	14	4	27	1	13	133	2
j	6	26	4	1	0	6	0	0	2	255

Prediction on Test Data-Random Forest

	a	b	c	d	e	f	g	h	i	j
a	144	0	25	12	22	14	37	17	23	6
b	0	265	0	0	0	23	0	0	0	12
c	32	12	74	12	10	58	42	29	18	13
d	28	2	21	50	15	16	15	42	60	1
e	5	0	0	2	222	0	5	16	0	0
f	0	18	37	1	0	121	0	10	4	9
g	48	0	40	3	11	11	92	39	5	1
h	4	0	27	7	14	17	48	176	7	0
i	18	17	18	27	4	38	3	15	107	3
j	1	105	19	0	0	35	4	1	10	125

Prediction on Test Data-SVM

Figure 3 A DT, k-NN, RF, and SVM algorithm error matrix displaying accurate and wrong cross-tabulations of evaluation sample. The LCLU classes are numbered "a" through "j", where a = Annual Crop, b = Forest, c = Herbaceous Vegetation, d = Highway, e = Industrial, f = Pasture, g = Permanent Crop, h = Residential, i= River, j= Sea Lake. The performance indicators presented in Table on Validate Data are derived from the confusion matrices.

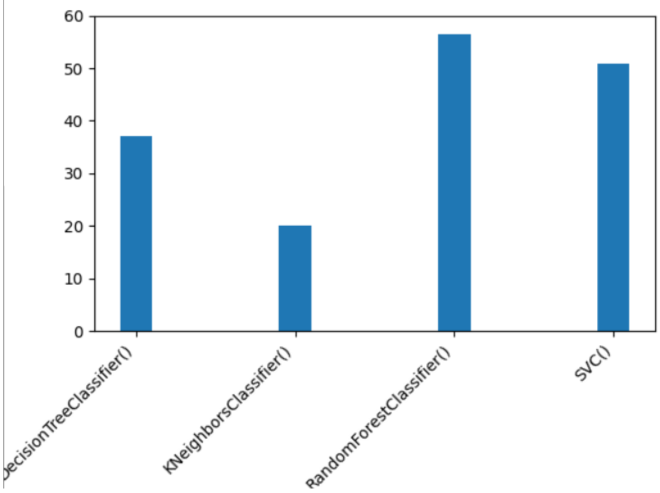


Figure 4- Graphical representation of DT, k-NN, RF and SVM-performance on Test data

Table 3. Validation results of classification algorithms

Validation Method	DT	k-NN	SVM	RF
Overall Accuracy (%)	37.22	21.25	50.85	56.70

	a	b	c	d	e	f	g	h	i	j
a	120	3	36	21	16	17	34	21	21	11
b	3	198	4	2	0	14	1	0	10	68
c	43	7	67	26	7	32	33	32	34	19
d	38	5	39	41	9	16	33	29	34	6
e	24	0	21	22	106	3	35	29	8	2
f	13	16	32	12	0	50	15	19	26	17
g	50	0	42	25	22	13	39	44	13	2
h	24	2	46	39	22	25	42	88	11	1
i	26	11	27	37	10	32	16	15	71	5
j	8	34	4	4	1	19	1	1	3	225

Prediction on Validate Data-Decision Tree

	a	b	c	d	e	f	g	h	i	j
a	187	1	23	1	0	11	0	0	4	73
b	0	0	0	0	0	0	0	0	0	300
c	50	1	37	0	0	7	0	0	3	202
d	67	8	49	15	0	20	0	0	7	84
e	184	0	24	2	0	2	1	0	4	33
f	3	0	15	0	0	9	0	0	0	173
g	97	0	26	1	0	6	0	0	0	120
h	33	2	18	0	0	4	0	0	0	243
i	63	12	33	0	0	25	0	0	34	83
j	4	0	2	0	0	2	0	0	0	292

Prediction on Validate Data-k-NN

	a	b	c	d	e	f	g	h	i	j
a	206	1	25	6	4	5	15	13	22	3
b	0	268	0	0	0	21	0	0	1	10
c	39	2	83	20	6	32	37	42	35	4
d	33	8	16	56	15	11	18	39	52	2
e	11	0	1	5	206	0	10	17	0	0
f	1	13	50	3	0	90	0	18	19	6
g	63	0	34	7	17	11	58	53	7	0
h	14	0	26	15	24	8	27	183	3	0
i	21	15	21	22	10	17	4	13	127	0
j	6	24	3	0	0	9	0	0	4	254

Prediction on Validate Data-Random Forest

	a	b	c	d	e	f	g	h	i	j
a	162	3	27	9	12	8	44	11	19	5
b	0	266	0	0	0	26	0	0	0	8
c	28	8	63	6	9	69	47	35	24	11
d	31	7	20	56	15	17	17	48	35	4
e	5	0	1	3	218	0	6	17	0	0
f	1	16	50	2	0	95	1	9	14	12
g	56	0	41	4	12	18	75	34	8	2
h	3	0	32	8	12	17	30	192	6	0
i	29	19	25	21	10	27	6	10	102	1
j	4	96	8	0	0	40	1	0	7	144

Prediction on Test Data-SVM

Figure 5 A DT, k-NN, RF, and SVM algorithm error matrix displaying accurate and wrong cross-tabulations of evaluation sample. The LCLU classes are numbered "a" through "j", where a = Annual Crop, b = Forest, c = Herbaceous Vegetation, d = Highway, e = Industrial, f = Pasture, g = Permanent Crop, h = Residential, i= River, j= Sea Lake. The performance indicators presented in Table on Validate Data are derived from the confusion matrices.

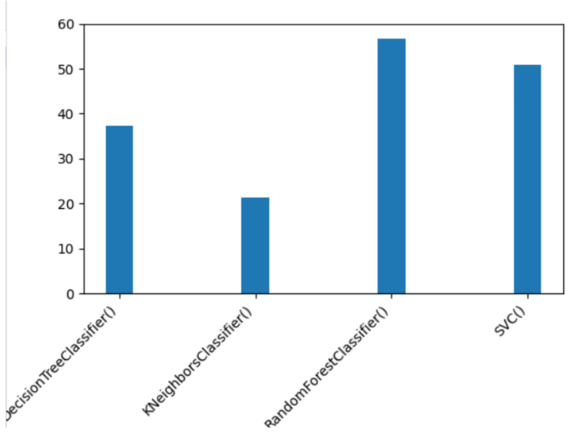


Figure 6- Graphical representation of DT, k-NN, RF and SVM -performance on validate data

The Random Forest approach appears to be superior to decision tree, k-nearest neighbor, and SVM in image classification, although SVM and decision tree perform similarly well, while k-NN performs worse than the other algorithms. When it comes to producer's accuracy, RF scores 94.4 percent in the sea lake class and SVM scores 74.49 percent in the industrial class, with 100% UA and 50.85 percent total accuracy for each class separately.

Conclusion

For land use and land cover image classification, RF is used in conjunction with SVM, decision tree and k-nearest neighbor methods in this research and in figure 3,4,5, and 6. For low-resolution satellite image classification, the results of a statistical analysis show that RF can be a superior alternative approach than other methods, especially when compared to other methods. The findings show that the SVM consumes a lot of computing time but has a high level of accuracy compared to other techniques, like DTs and k-Nearest Neighbor. It is easier to assess the performance of DTs with smaller data sets than it is with larger ones. If you want to use the Random Forest algorithm, you will need a lot of resources and computational power. It also requires a considerable amount of time to learn as it relies on several decision trees to make its classification.

Authors' contributions

Reena Thakur is a research scholar of Medi-Caps University, Indore, Madhya Pradesh, India. She has developed the techniques for removing speckle noise from satellite images, responsible for providing idea, designing model and experimental methods under the guidance of Dr. Prashant Panse, Associate Professor, Medi-Caps University, Indore, Madhya Pradesh, India.

References

[1]. S. E. Franklin "Remote Sensing for Sustainable Forest Management". New York. Lewis Publishers, 2001.

[2]. A. De Sherbinin "Land use and land cover change" in 4 A CLESIN Thematic Guides, Palisades, YN. Centre

for International Earth Science Information Network of Columbia University. 2002, pp. 67.

[3]. Dean, A.M and Smith, G.M. "An evaluation of per – parcel land cover mapping using maximum likelihood class probabilities" in International Journal of Remote Sensing, 24 (14), 2905-2920, 2003.

[4]. EuroSAT: A novel dataset and deep learning benchmark for land use and land cover classification. Patrick Helber, Benjamin Bischke, Andreas Dengel, Damian Borth. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019.

[5]. Introducing EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. Patrick Helber, Benjamin Bischke, Andreas Dengel. 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018.

[6]. Anusha, D. J. ., R. . Anandan, and P. V. . Krishna. "Modified Context Aware Middleware Architecture for Precision Agriculture". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 7, July 2022, pp. 112-20, doi:10.17762/ijritcc.v10i7.5635.

[7]. M. Seetha, I. V. Muralikrishna "Artificial neural network and other methods of image classification" in Journal of theoretical and Applied Information Technology.

[8]. Geospatial Innovation Facilities in Center for Biodiversity and Conservation.

[9]. Cortes, C., and V. Vapnik. 1995. "Support-vector Networks." Machine Learning 20 (3): 273–297. doi:10.1007/bf00994018.

[10]. Vapnik, V. 1982. Estimation of Dependences Based on Empirical Data: Springer Series in Statistics (springer Series in Statistics). New York, NY: Springer-Verlag.

[11]. Mountrakis, G., I. Jungho, and C. Ogole. 2011. "Support Vector Machines in Remote Sensing: A Review." ISPRS Journal of Photogrammetry and Remote Sensing 66 (3): 247–259. doi:10.1016/j.isprsjprs.2010.11.001.

[12]. Ravita, R., & Rathi, S. (2022). Inductive Learning Approach in Job Recommendation. International Journal of Intelligent Systems and Applications in Engineering, 10(2), 242–251. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/1829

[13]. Breiman, L. 2001. "Random Forests." Machine Learning 45 (1): 5–32. doi:10.1023/a:1010933404324.

[14]. Hastie, T., R. Tibshirani, and J. Friedman. 2009. "Random Forests." In The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 587–604. New York: Springer New York.

[15]. Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 46(3), pp.175-185.

[16]. Quinlan, J.R., 1986. Induction of decision trees. Machine learning, 1(1), pp.81-106.

[17]. Staff, I.,2017. Decision Tree. Investopedia. Retrieved 7 November 2017, from http://www.investopedia.com/terms/d/decision-tree.asp

[18]. Woznicki, S. A., J. Baynes, S. Panlasigui, M. Mehaffey, and A. Neale. 2019. "Development of a Spatially Complete Floodplain Map of the Conterminous United States Using Random Forest." Science of the Total Environment 647: 942–953. doi:10.1016/j.scitotenv.2018.07.353.

[19]. Betts, M. G., W. J. Christopher Wolf, B. P. Ripple, K. A. Millers, S. H. Adam Duarte, M. Butchart, and T. Levi. 2017. "Global Forest Loss Disproportionately Erodes Biodiversity in Intact Landscapes." Nature 547: 441. doi:10.1038/nature23285.

[20]. Araki, S., M. Shima, and K. Yamamoto. 2018. "Spatiotemporal Land Use Random Forest Model for Estimating Metropolitan NO2 Exposure in Japan." Science of the Total Environment 634: 1269–1277. doi:10.1016/j. scitotenv.2018.03.324.

[21]. Rodriguez-Galiano, V. F., B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez. 2012. "An Assessment of the Effectiveness of a Random Forest Classifier for Land-cover Classification." ISPRS Journal of Photogrammetry and Remote Sensing 67: 93–104. doi:10.1016/j.isprsjprs.2011.11.002.

[22]. Gupta, D. J. . (2022). A Study on Various Cloud Computing Technologies, Implementation Process, Categories and Application Use in Organisation. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(1), 09–12. https://doi.org/10.17762/ijfrcsce.v8i1.2064

[23]. Wu, L.; Zhu, X.; Lawes, R.; Dunkerley, D.; Zhang, H. Comparison of machine learning algorithms for classification of LiDAR points for characterization of canola canopy structure. *Int. J. Remote Sens.* **2019**, *40*, 5973–5991.

[24]. Aaron E. Maxwell, Timothy A. Warner & Fang Fang (2018) Implementation of machine-learning classification in remote sensing: an applied review, International Journal of Remote Sensing, 39:9, 2784-2817, DOI: 10.1080/01431161.2018.1433343

[25]. Geospatial Innovation Facilities in Center for Biodiversity and Conservation.

[26]. Sally Fouad Shady. (2021). Approaches to Teaching a Biomaterials Laboratory Course Online. Journal of Online Engineering Education, 12(1), 01–05. Retrieved from http://onlineengineeringeducation.com/index.php/joee/article/view/43

[27]. Sultan AlSultan "Innovative Satellite Image Map Of R. Alkhabrta Area, Saudi Arabia Using High Resolution Image", School of Physics.