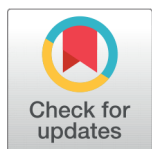


## RESEARCH ARTICLE



### OPEN ACCESS

**Received:** 07-05-2020

**Accepted:** 14-05-2020

**Published:** 15-06-2020

**Editor:** Dr. Natarajan Gajendran

**Citation:** Junejo KN (2020) A novel pixel-based supervised hybrid approach for prediction of land cover from satellite imagery. Indian Journal of Science and Technology 13(17): 1786-1794. <https://doi.org/10.17485/IJST/V13i17.538>

#### \*Corresponding author.

Khurum Nazir Junejo

Program Architect, Ibex Global, Centre Point, Shaheed e Millat Expressway, Karachi, Pakistan. Tel.: +92-333-308-2902 [junejo@gmail.com](mailto:junejo@gmail.com)

**Funding:** None

**Competing Interests:** None

**Copyright:** © 2020 Junejo. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](https://www.indst.org/))

## A novel pixel-based supervised hybrid approach for prediction of land cover from satellite imagery

Khurum Nazir Junejo<sup>1\*</sup>

<sup>1</sup> Program Architect, Ibex Global, Centre Point, Shaheed e Millat Expressway, Karachi, Pakistan. Tel.: +92-333-308-2902

### Abstract

**Background/Objectives:** To determine the land use/cover from satellite imagery using image enhancement, image processing, and supervised machine learning techniques. This land usage will help in land use policy development, disaster assessment, planning of urban infrastructure, forest and agriculture monitoring and conservation. **Methods/Statistical analysis:** A pixel-based supervised hybrid machine learning approach is used that combines parametric density estimation followed by a k-nearest neighbor (k-NN) classifier to predict whether a particular pixel belongs to a nucleated village or to a field, forest, river, or some other terrain from satellite images. Spatial and texture features derived from the images are used as features for the models. Free satellite images crawled from Google Earth<sup>TM</sup> are hand labeled to serve as the ground truth for training and testing Models. Models are evaluated using four-fold cross validation. Comparison with other related techniques is also presented. **Findings:** Our experiments suggest that instead of using pixel intensity values as features, the intensity values after edge detection give better prediction accuracy. The parametric density estimation for the two classes is better modeled as Rayleigh distribution than as a normal distribution. Smoothing further helps improve the accuracy of the models. Passing the prediction of the parametric density estimator classifier through k-NN further reduces the error by removing the salt-and-pepper effects. Effects of using different size and number of smoothing filters is also discussed. Additionally, different parameters for k-NN were also evaluated to find the best models. The model was successful in achieving a very high accuracy of more than 97% with a very small false positive rate. The results demonstrate the successful discrimination of land cover of towns and nucleated villages from the surrounding terrain. **Novelty/Applications:** A pixel-based hybrid approach is proposed to improve accuracy of land cover classification. The proposed features and image enhancement techniques also help improve the prediction significantly. Using the proposed technique, the covered area of nucleated villages/towns can be determined for assessing the growth of towns/villages, for better of planning of urban infrastructure and land use policy, for disaster assessment, and forest and agriculture monitoring and conservation.

**Keywords:** Land cover classification; land usage classification; satellite image classification; nucleated village segmentation

# 1 Introduction

Making of land cover and usage maps is as old as history itself. Traditionally they were made with the help of scouts and explorers. It is only recently that these maps are being made through aerial and satellite images. The earliest interpretation of aerial images for land cover classification started in 1950s<sup>(1,2)</sup>. It was in 1972 that Earth Resource Technology Satellite which is now known as Landsat was launched into orbit<sup>(3)</sup>. Landsat opened the door of remote sensing using satellite imagery for land cover classification. Although a commercial resource at first, now Landsat provides free satellite images for land cover and usage analysis. More recently in 2015 and 2017, European Space Agency launched Sentinel-2 A and Sentinel-2 B satellites with better remote sensing capabilities<sup>(4)</sup>. These satellites along with the satellite imagery freely provided by Google Earth<sup>TM</sup> have resulted in hundreds of studies that have developed land cover type maps that are being used in numerous applications such as land use policy development, crop mapping, ecosystem services, infrastructure planning, disaster assessment, conservation, forest management, agricultural monitoring, sustainable development, nature protection, and dynamic assessment of land use/cover<sup>(5–7)</sup>. For example, <sup>(8)</sup> show in Figure 1. How deforestation has occurred at a place near Villamontes, Bolivia.

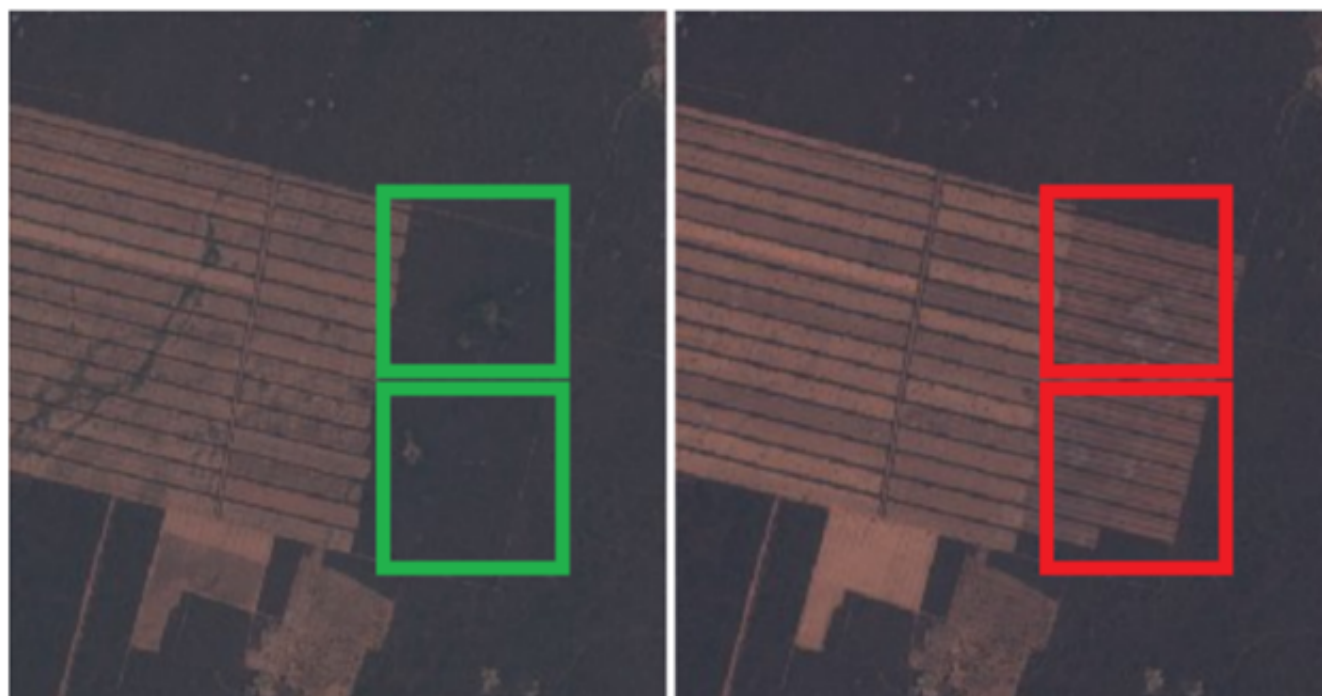


Fig 1. Deforestation Detection Villanmontes, Bolivia. Left and right image taken in 2015 and 2016, respectively. Image courtesy: <sup>(8)</sup>

Of all these various applications we focus on geo-spatial data of nucleated villages and towns because their locations, size, population, and other parameters can greatly help decision makers in many developing countries. For example, it can help in determining the area occupied by the village, and hence its growth over the years. In case of a disaster, such as a flood, it can be used to estimate the extent of the inundated area. We employ twelve different images having different types and sizes of nucleated villages and the surrounding terrain to train and test our model. After applying the necessary pre-processing, feature engineering, and image enhancements we are able to accurately predict which area on the image belongs to the village. For model training and parameter tuning we have used four-fold cross validation. The model consists of a parametric model followed by a k-Nearest Neighbor (k-NN) algorithm. A very high accuracy of more than 97% is achieved with a low False Positive (FP) rate. The effects of varying the parameter values and the image enhancement techniques are also discussed in detail.

The rest of the paper is organized as follows: in Section 2 we provide the related work, followed by the materials and methods in Section 3. The results along with the discussion is provided in Section 4. We finally conclude in Section 5.

# 2 Related Work

In the early 1970s, satellite images were manually classified into various land cover types through the examination of printed aerial images. The images were printed as black and white composite or individual bands<sup>(3)</sup>. Later, the classification was done through knowledge-based systems in which the pre-defined rules were programmed into software products that classified land areas into different cover type<sup>(9,10)</sup>. With the start of this millennium machine learning (ML) and pattern recognition (PR) systems have emerged as the leading approach for satellite image classification<sup>(11,12)</sup>. ML and PR methods alleviate the need of a domain expert who is responsible for coming up with rules to classify images. Furthermore, using ML and PR better accuracies are achieved as these methods can find hidden and complex patterns that are beyond the comprehension of the domain experts<sup>(13)</sup>. ML and PR methods can be either unsupervised or supervised. Unsupervised techniques do not require human labeled images for model training and testing, whereas supervised approaches require parts of an image to be labeled into the different land type categories that are of interest to the end user. Our approach belongs to this latter category.

Depending on the resolution of the images, land cover classification technique can be either pixel, sub-pixel, patch, or object based. In a pixel-based approach, it is assumed that a pixel is made up of one homogenous land cover type therefore each individual pixel is assigned a category<sup>(14,15)</sup>. However, for low resolution images many pixels are composed of more than one land cover type. In such scenarios sub-pixel-based approaches such as fuzzy set approaches

are used<sup>(16)</sup>. A patch-based approach considers classification based on a window of a specific size e.g. a 8 x 8 size pixel window that consists of 64 pixels may be classified as belonging to a particular class<sup>(8)</sup>. The object-based approach on the other hand classifies images using a grouping or clustering of pixels<sup>(14,17)</sup>.

Various techniques for Land cover classification have different number and types of categories. For example<sup>(18)</sup> use satellite imagery data for retrieving leaf area index and leaf and canopy chlorophyll content of potato crops. Whereas<sup>(19)</sup> classifies the images into six different land cover types, namely: barren land, grassland, road, trees, grassland, water bodies and buildings. We on the other hand classify the land cover into two types or classes, namely village/town and terrain. Our work is perhaps most similar to<sup>(12)</sup> having used the same land cover types and source satellites but differ in approach and features used.

Multi-temporal spatial images derived from a time series of satellite images are used to assess the changes over time<sup>(16,20)</sup>. Our approach is time independent but can be applied on images from different timelines to determine the difference between the two images over time. This study is also not about using dimensionality reduction<sup>(21)</sup> or removing cloud cover from images<sup>(22)</sup> or other image cleaning and pre-processing.

Several approaches try to make use of images from multiple new and advanced remote sensors such as advanced very high-resolution radiometer (AVHRR), radio detection and ranging (Radar), moderate-resolution imaging spectroradiometer (MODIS), synthetic aperture radar (SAR), light detection and ranging (LiDAR), and panchromatic images<sup>(23,24)</sup>. Different signals from different sensors are then fused together to increase the accuracy by complementing each other. Our approach does not make use of any of these advanced sensor data and instead just relies on imagery provided by Google Earth<sup>TM</sup> and uses only the pixel color intensity values. This makes our approach easier, cheaper, and more scalable.

## 3 Materials and Methods

### 3.1 Data Collection and Preparation

#### 3.1.1 Data Collection

The data for this study was collected from free satellite imagery resource Google Earth<sup>TM</sup>. Various images were captured but only 12 images were selected that had villages/town surrounded by agricultural fields or barren land, or rivers or any other terrain. We chose only 12 images because it is very time consuming to manually hand label these images for ground truth. Each image has dimensions of 2048 by 2048 pixels. Figure 2 shows two of the twelve images used for this study. What makes this challenging is the variation introduced by different land cover types, different sizes and shapes of villages, and the variation introduced by the fact that Google Earth<sup>TM</sup> captures images at different times from different sensors under varying atmospheric conditions. We hand labeled all the images by demarcating the village boundary to make the ground truth. We use this ground truth for training and evaluation of our models. The challenge is labelling these images was that it was unclear what the boundaries of the villages are as there are spread far wide with many individual buildings surrounded with terrain. Therefore, we have labelled houses as belonging to the village, but the adjoining open spaces and farmland are considered as part of the terrain. Only the pixel values of the images such as in (a) of Figure 2 were used, and no other data from advanced remote sensors such as AVHRR, LiDar, MODIS, etc. was collected or used.

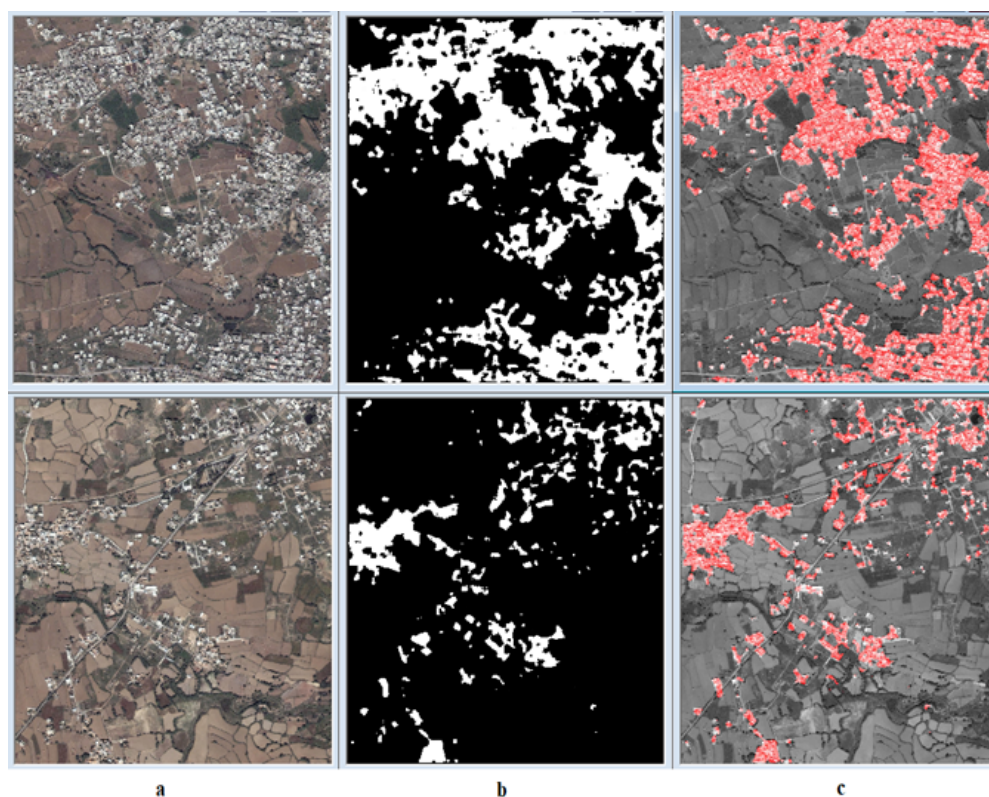


Fig 2. Two Sample Images from Training Data and their Ground Truth. (a) Original Image, (b) Hand Labelled Ground Truth, and (c) Ground Truth Super Imposed on the Original Image



### 3.1.2 Data Split for Training and Evaluation

To train and then test the model we need to split the data into two sets, namely, the training and test set. We split the data between the training and test by randomly choosing 9 of the 12 images for training and the rest for testing. This process is repeated four times; thus, each iteration takes different training and testing images, a technique that is referred to in the machine learning literature as cross-validation. The more iterations that we perform the better the estimate of the classifier's true performance, but this would lead to very long-running time of the algorithm, therefore we have restricted the experiments to four iterations only. Hence, all the results reported in this manuscript, the tables, etc. are an average of the above mentioned four iterations for the same parameters and models, until and unless specified otherwise.

## 3.2 Feature Engineering

Looking at the images of the houses and fields, one can notice that the variations of texture where there are houses is much more than that of the fields, that is because the texture of the fields is quite smooth. From this observation, it can be inferred that if we were to do edge detection on these images, the areas where there are houses will have more edges than the areas with fields. The intuition behind this observation is that there is larger texture variation in houses, and also because of the small size of the houses, there are many edges close by as compared to that in the fields, because a field block is almost equivalent in size to 10 to 20 houses. Therefore, there would only be edges running along the perimeter of the field block.

### 3.2.1 Distribution Analysis and Edge Detection

For determining the validity of this hypothesis, we synthesized two ground truth images, one for houses and one for fields. These two ground truths are shown in Figure 3. Many different images were sampled and patched together to make this true representative image. Edge detection (ED) was then performed and the histogram of the intensity values of the two images before and after ED is plotted in Figure 4. The difference in the histograms before and after ED suggests that the amount of overlap has decreased but even then, there is quite a significant overlap between the two distributions suggesting that this feature is not good enough for discriminating the classes.

### 3.2.2 Smoothing and Image Enhancement and Smoothing

Based on the observation mentioned earlier in this section, the edge density in areas having houses is very high as compared to that of the fields. Therefore, it is quite intuitive to use smoothing (or blurring) on the synthesized images because the edges in the fields are surrounded with more pixels with zero or close to zero intensity values as compared to that of the houses because houses are very close by thus forming a cluster. As a result, the intensity values of edges in the fields will move more closer towards zero as compared to edges in the areas containing houses. Figure 4. shows that this observation is correct. It can be seen from this figure that as we increase the number of times we smooth the data the distribution of the fields moves closer and closer to left side of the intensity scale. Whereas, the distribution of the houses started off with negative exponential distribution (Figure 3) to Rayleigh distribution (Figure 4) as the smoothing is increased to five times after ED, and then finally to a normal distribution when smoothing is done 10 times after ED.



Fig 3. Synthesized Images for Ground Truth (a) Images of various villages patched together (b) images of various terrain patched together

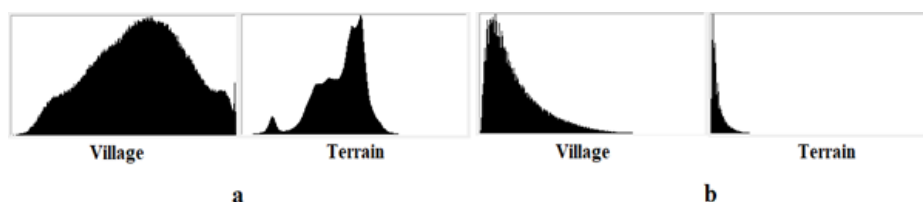
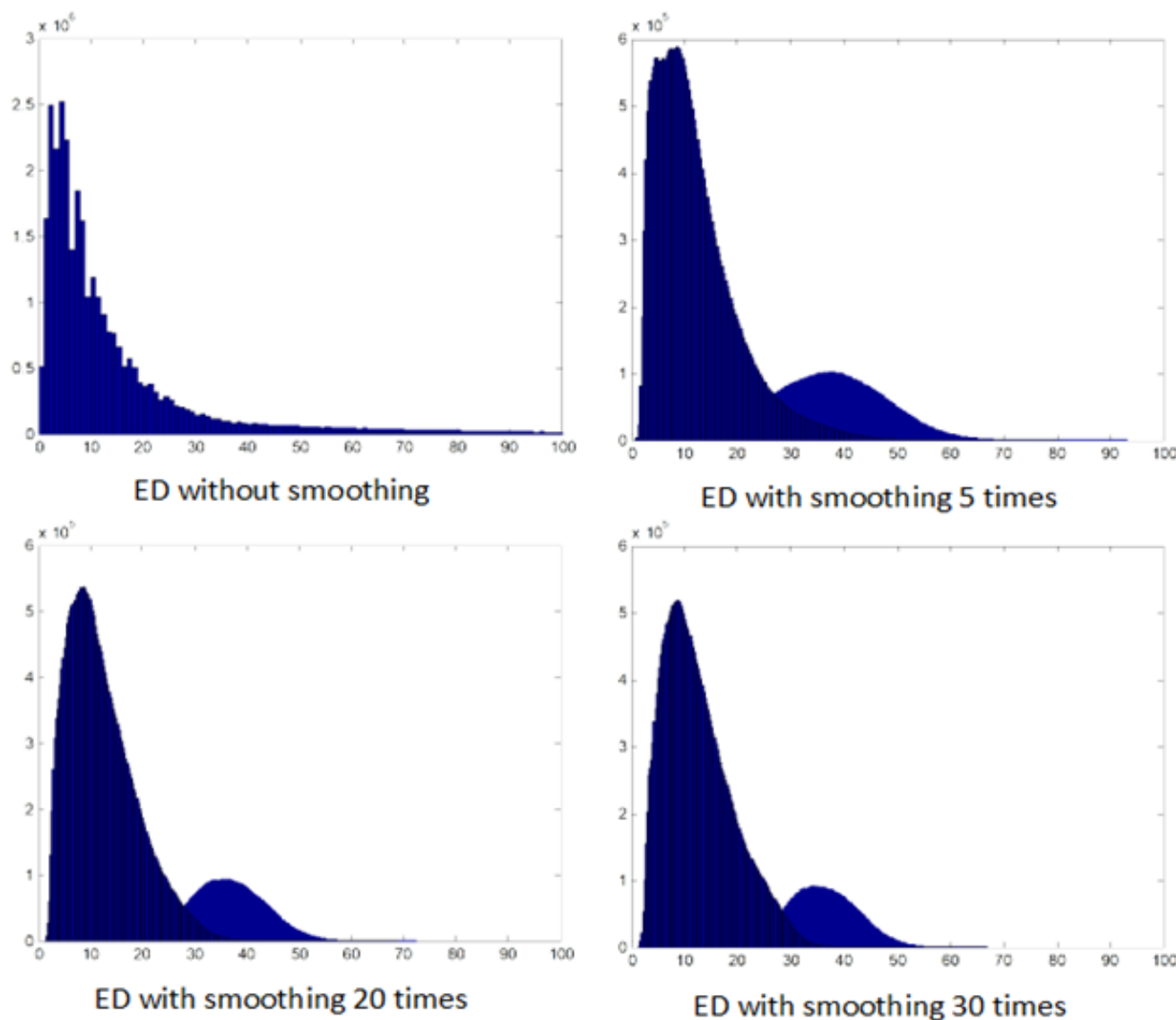


Fig 4. Histogram of Pixel Intensity Values of the Two. Classes Histogram of intensity values of images in Figure 3. (a) before edge detection (ED) (b) after ED

Similar trend is seen for our 9 training images, the average intensity values for which are given in Figure 5. It is obvious from this figure that before smoothing the distribution of both the classes there was almost a complete overlap between each other but as we increase the number of smoothing the two distributions

start to split apart. There is also a limit to how many times we can smooth the images. As smoothing is done many times (30 times) we can see that the advantage getting lesser as a fair amount of values of the distribution of fields now lies under the distribution of the houses because after smoothing the intensity values of field pixels surrounding the house pixels also increases. From this figure it can be further inferred that as we increase smoothing, the distribution of houses more and more resembles that of a normal distribution while the distribution of fields resembles Rayleigh distribution.



**Fig 5.** Distribution Discrimination Using Image Smoothing. Peakier distribution curve corresponds to the pixels belonging to the village whereas the flatter distribution curve corresponds to the pixels belonging to the terrain

### 3.3 Parametric Density Estimation

The distribution of houses resembles that of a normal distribution while the distribution of fields resembles Rayleigh distribution as seen in Figure 5. We, therefore, model the two classes through parametric density estimation. The feature vector is one dimensional having 256 distinct values for the 256 unique pixel intensity values. Therefore, approximately 256 probabilities for the village class ( $P(\text{pixel}/\text{village})$ ) and 256 for the terrain class ( $P(\text{pixel}/\text{terrain})$ ), need to be estimated from the nine training images which totals to around 37 ( $2048 \times 2048 \times 9$ ) million pixels. Given such a huge number of pixels to estimate our probabilities the parametric density approach adopted here even though being quite simple give very good results. Prior probabilities are not taken into account because in some images the area covered by the village very high while in some it is the opposite. Since there is no fixed percentage of houses or fields in an image, therefore, the priors of the two distributions is not used in modeling.

Initially both the distributions were thought to be best modelled by Rayleigh distribution (Figure 5). Further experiments showed that the village class would be better modeled as a normal distribution. This was realized after the densities of the two classes were modeled as both as Rayleigh and as normal distribution as shown in Figure 6. The next thing is to find the decision boundary. After having determined the decision boundary we are left with a very simple

decision rule to classify between the village and the terrain class which is as follows:

if the value of the feature is greater than the decision boundary then classify it as belonging to village otherwise classify as terrain.

### 3.3.1 Determining Decision Boundary

The Rayleigh distribution could be given as

$$f_1(x) = \frac{xe^{\frac{-x^2}{2\sigma_1^2}}}{\sigma_1^2} \quad (1)$$

to find the decision boundary when both the distributions are Rayleigh, we equate  $f_1(x) = f_2(x)$ . Solving this equation, we get the following expression for the decision boundary of the distributions:

$$x = \sqrt{\frac{\log(\sigma_1^2 - \sigma_2^2) * \sigma_1^2 * \sigma_2^2}{\sigma_1^2 - \sigma_2^2}} \quad (2)$$

The MLE of  $\sigma^2$  can be given as

$$\hat{\sigma}^2 = \frac{1}{2N} \sum_{i=1}^n x_i^2 \quad (3)$$

The normal distribution could be given as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)$$

To find the decision boundary of this case when one distribution is normal and the other distribution is Rayleigh we solve for x the following equation numerically

$$\frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}} - \frac{xe^{\frac{-x^2}{2\sigma_1^2}}}{\sigma_1^2} = 0 \quad (5)$$

## 3.4 K-NN

The result of the parametric density classifier is then fed as input to k-NN classifier to get the final prediction. The rationale for this is the observation that the images classified by the parametric model have regions within the village area that were classified as terrain and similarly there were pixels within the terrain that were classified as village pixels. To remove this salt-and-pepper effect in prediction or in other words to remove these types of False Positive (FP) and False Negatives (FN) errors k-NN classifier is used. The results also demonstrate improvement with the use of k-NN.

In this problem setting k-NN degrades to just counting the number of pixels that belong to both classes within a spatial window of k pixels and then assigning the class which has the highest number of pixels in that window. Furthermore, the closest neighbors here would be the spatial neighbors and not the neighbors closest in terms of the value of the feature. Along the borders of the image, zero padding could change the results significantly. It is therefore avoided by replicating the value along the border. Different window sizes for the k-NN classifier are used it: 1x1 (same as density estimation without the k-NN), 3x3, 5x5 and 9x9.

## 3.5 Smoothing Parameters

Following are three different parameters regarding the smoothing filter applied to the images:

- 1- Size of the smoothing filter
- 2- Number of times the filter should be applied
- 3- Gaussian or uniform blurring filter

There are various tradeoffs between selecting different parameters for smoothing. For example the advantage of selecting a large filter is that since the area surrounding the edges of fields is dark so these edges will blur quickly. On the other hand the drawback with this is that homes that are not located in a clutter or the homes on the outer rim of the clutter will also get blurred to black. If a filter that is too small is used, then edges of field will not get blurred because they have a small neighborhood of high intensity values. The number of times to smooth is also crucial because if you smooth only few times the discriminability between the classes will not increase significantly, but on the other hand if you apply the filter too many times then it would also not help because eventually the intensity of each pixel will converge to the mean intensity value of the image. Gaussian filter weighs the contribution of each pixel in accordance with its proximity while the uniform filter will give equal weightage to each pixel in the filter window. There a bigger window might be required for Gaussian smoothing, furthermore it might have to be applied greater number of times than that of a uniform smoothing filter thus taking more time.

## 4 Results and Discussion

### 4.1 Result of Parametric Density Estimation

Table 1 shows the results of the two different parametric models. Rayleigh-Rayleigh model (RRM) refers to the model in which the distribution of the both the classes is modeled using a Rayleigh distribution, whereas, Rayleigh-Normal (RRN) refers to the model in which the terrain class is modeled as a Rayleigh distribution but the village class is modeled as a normal distribution. Surprisingly, even though Rayleigh-Rayleigh model (RRM) does not fit the distribution as well as Rayleigh-Normal model (RNM) does (Figure 6), RRM performs better than RNM. RRM achieves both, the overall minimum error and as well as the least error for a given number of smoothing. This result is counter intuitive because both the training and test data undergo the same transformation and in the transformed space RNM better approximates the distribution (Figure 6).

Table 1. Result of parametric density estimation. Error, FP and TP rate values are in percentages.

# of times Smoothing	Rayleigh-Rayleigh Model			Rayleigh-Normal Model		
	Error	FP rate	TP rate	Error	FP rate	TP rate
0	15.19	5.05	39.49	18.00	8.23	48.44
5	6.19	4.20	82.40	7.69	7.79	92.08
10	5.46	4.28	89.68	6.06	6.55	94.83
15	4.16	3.35	89.37	5.73	6.45	97.17
20	3.47	2.43	88.32	5.27	6.19	97.89
25	3.82	2.72	90.85	5.57	6.11	96.69
30	4.79	4.58	93.26	5.06	5.46	94.93
35	3.87	2.67	90.26	5.53	6.20	96.49
40	5.15	4.56	92.49	5.58	6.18	96.97

One thing to notice though is that RNM has higher TP rate than that of RRM which is a good thing, but at the same time it has a higher FP rate than RRM, the reason for this is evident from Figure 6. In this figure it can be seen that the decision boundary of RRM lies somewhere around 21 on the x-axis while the decision boundary for the RNM lies somewhere around 24 x-axis, because of which RNM model misclassifies less of the village pixels as fields pixel as a result higher TP rate and as a consequence higher FP rate as well. In a scenario where the cost of FP is higher then RNM is better suited for classification.

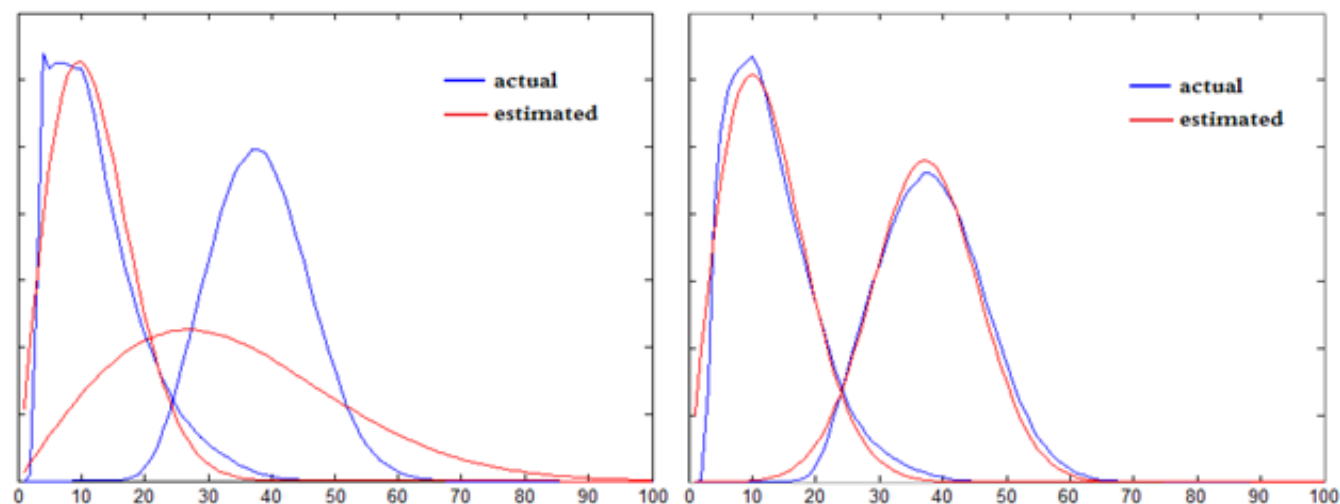


Fig 6. Parametric Density Estimation On the left both the densities estimated as Rayleigh distribution whereas, on the right, terrain class modeled as Rayleigh distribution and village class modelled as normal distribution. The densities were estimated after edge detection followed by smoothing 10 times.

### 4.2 Effects of Applying k-NN

Table 2 shows the significant improvement in the error rate with the application of k-NN. One can argue that the same effect can be obtained with more smoothing, but our results show otherwise. We can also notice that with a bigger neighborhood size i.e. 15x15 we get a better performance gain than a smaller one like 1x1. Results of Table 3 demonstrate that applying k-NN multiple times helps reduce the error, it will help to do that until certain number of times after which the performance will start to degrade. We cannot verify this claim because applying the k-NN filter is quite expensive and time-consuming because we perform 5 iterations over it to give the averaged result as well.

**Table 2.** Result of Applying k-NN to Output of Density Estimators. Error, FP and TP rate values are in percentages.

Neighborhood Size	Smoothing 5 Times			Smoothing 20 Times		
	Error	FP rate	TP rate	Error	FP rate	TP rate
1x1	6.93	4.68	82.18	4.87	3.57	88.85
15x15	5.78	3.76	84.45	4.76	3.38	88.57

**Table 3.** Effect of applying k-NN multiple times. Error, FP and TP rate values are in percentages. The results are for the best parameters with smoothing done 20 times.

Neighborhood Size	Smoothing 5 Times		
	Error	FP rate	TP rate
1x1	2.68	1.57	86.49
9x9 1 time	2.64	1.53	86.48
9x9 5 times	2.56	1.39	86.01

### 4.3 Smoothing Parameters

Experiments with different parameters of the smoothing filter were carried out to obtain the best set of features. For a fixed set of training and test examples, results of applying different filter sizes is given in Table 4. It shows that with the increase in the filter size, the error rate decreases until a certain filter size (15x15) and after which it starts to increase. However, the optimal feature can not only be determined by the filter size but the number of times we apply that filter is also important. For smaller filter sizes we will have to smooth many times more in contrast to if we were using a bigger filter size. The results in Table 4 are obtained after doing smoothing 10 times, but if you do smoothing 20 times you get an even better performance with a filter of size 7x7 i.e. around 3.47% error rate (refer to Table 1). Similarly, if you perform smoothing only 3 times you get around 4.76% error rate (Table 5). Therefore, we conclude that for smaller filter size we need to perform smoothing more number of times than for larger filter sizes. A question arises then, why not always choose a very big filter. The answer to this question is that with a big filter we get more error (see Table 5) because the houses that are located at the corners of the town or are surrounded by fields will get blurred out. Then one might think that we should choose the smallest filter size, the problem with this argument is that if you choose it to be 1x1 or 2x2, there will be no significant effect. Filter size of 3x3 will benefit the performance of the classifier but then you will need to apply it many times and the cost will become restrictive. As a tradeoff between efficiency and performance we recommend a filter size between 5x5 to 9x9.

**Table 4.** Experimenting with different blurring filter sizes. All results reported on a fixed set of mutually exclusive training and testing data after a single run. Number of times smoothing done is 10 times for all runs. Value of Error, FP and TP rates are in percentage values.

Blurring filter size	Error	FP Rate	TP Rate
3x3	9.24	5.13	70.92
7x7	5.67	4.19	87.02
11x11	4.75	3.36	88.49
15x15	5	2.76	84.22
19x19	5.95	2.42	77.02

**Table 5.** Amount of smoothing required to reach the optimal results with the filter of size 19x19

# of Times Smoothing	Error	FP Rate	TP Rate
2	5.01	3.80	89.13
3	4.76	3.44	88.85
5	4.82	2.97	86.25
7	5.18	2.67	82.69
10	5.95	2.42	77.02

### 4.4 Comparison with Other Results

The performance of our model is comparable or better than the performance of models reported in other studies for similar settings. Our reported accuracy is better than all the pixel-based approaches in <sup>(25)</sup> and in Table 2 of <sup>(26)</sup>. The number of classes and the size of the training data varies though from the approaches in <sup>(26)</sup> and <sup>(25)</sup>.

## 5 Conclusion

Automatic classification of satellite imagery in different land cover types can help land use policy development, disaster assessment, planning of urban infrastructure, forest and agriculture monitoring and conservation. We propose a pixel-based ensemble approach to improve accuracy of land cover classification. The proposed hybrid approach combines parametric density estimation followed by a k-nearest neighbor (k-NN) classifier to predict whether a pixel belongs to



a nucleated village or to a field, forest, river, or some other terrain. The proposed features and image enhancement techniques also help significantly improve the prediction accuracy. Using the proposed technique, the covered area of nucleated villages/towns can be determined for assessing the growth of towns/villages, for better planning and conservation. We achieved an accuracy of more than 97% with very little FP rate. Results for different parameter selection choices and their discussion is also presented.

## Acknowledgement

I would like to thank Computer Vision Lab of Lahore University of Management Sciences for helping in acquiring the dataset. I would also like to thank Dr. Kamran Nishat of University of Houston for his discussions and useful insights.

## References

- Colwell RN. The photo interpretation picture in 1960. *Photogrammetria*. 1959. Available from: [https://doi.org/10.1016/0031-8663\(59\)80076-4](https://doi.org/10.1016/0031-8663(59)80076-4).
- Spurr SH. Aerial photographs in forest management. *Photogrammetria*. 1952. Available from: [https://doi.org/10.1016/S0031-8663\(52\)80004-3](https://doi.org/10.1016/S0031-8663(52)80004-3).
- Haack BN. Landsat: A tool for development. *World Development*. 1982. Available from: [https://doi.org/10.1016/0305-750X\(82\)90064-X](https://doi.org/10.1016/0305-750X(82)90064-X).
- Martins V, Barbosa C, de Carvalho L, Jorge D, Lobo F, Novo E. Assessment of Atmospheric Correction Methods for Sentinel-2 MSI Images Applied to Amazon Floodplain Lakes. *Remote Sensing*. 2017;9(4):322–322. Available from: <https://doi.org/10.3390/rs9040322>.
- Delalay M, Tiwari V, Ziegler AD, Gopal V, Passy P. Land-use and land-cover classification using Sentinel-2 data and machine-learning algorithms: operational method and its implementation for a mountainous area of Nepal. *Journal of Applied Remote Sensing*. 2019;13(1). Available from: <https://doi.org/10.1117/1.JRS.13.014530>.
- Zhu Z, Zhang J, Yang Z, Aljaddani AH, Cohen WB, Qiu S, et al. Continuous monitoring of land disturbance based on Landsat time series. *Remote Sensing of Environment*. 2020. Available from: <https://doi.org/10.1016/j.rse.2019.03.009>.
- Arévalo P, Olofsson P, Woodcock CE. Continuous monitoring of land change activities and post-disturbance dynamics from Landsat time series: A test methodology for REDD+ reporting. *Remote Sensing of Environment*. 2020. Available from: <https://doi.org/10.1016/j.rse.2019.01.013>.
- Helber P, Bischke B, Dengel A, Borth D. EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2019;12(7):2217–2226. doi:10.1109/jstars.2019.2918242.
- Taylor AM, Corr DG, Cross A, Hogg DC, Lawrence DH, Mason DC, et al. Development of a knowledge-based segmentor for remotely sensed images. *Philosophical Transactions of the Royal Society of London Series A Mathematical and Physical Sciences*. 1988;324(1579):437–483. Available from: <https://doi.org/10.1098/rsta.1988.0031>.
- Sikder IU. Knowledge-based spatial decision support systems: An assessment of environmental adaptability of crops. *Expert Systems with Applications*. 2009;36(3):5341–5347. doi:10.1016/j.eswa.2008.06.128.
- Moorthi SM, Misra I, Kaur R, Darji NP, Ramakrishnan R. Kernel based learning approach for satellite image classification using support vector machine. *2011 IEEE Recent Advances in Intelligent Computational Systems*. 2011;p. 107–110. Available from: <https://doi.org/10.1109/RAICS.2011.6069282>.
- Murtaza K, Khan S, Rajpoot NM. VillageFinder: Segmentation of Nucleated Villages in Satellite Imagery. In: and others, editor. *Proceedings of the British Machine Vision Conference*. 2009;p. 1–11. Available from: <https://doi.org/10.5244/C.23.83>.
- Abburu S, Golla SB. Satellite Image Classification Methods and Techniques: A Review. *International Journal of Computer Applications*. 2015;119(8):20–25. Available from: <https://doi.org/10.5120/21088-3779>.
- Yan G, Mas JF, Maathuis BHP, Xiangmin Z, Dijk PMV. Comparison of pixel-based and object-oriented image classification approaches—a case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing*. 2006;27(18):4039–4055. Available from: <https://dx.doi.org/10.1080/0143160600702632>.
- Kiani A, Ahmadi FF, Ebadi H. Developing an Interpretation System for High-Resolution Remotely Sensed Images Based on Hybrid Decision-Making Process in a Multi-scale Manner. *Journal of the Indian Society of Remote Sensing*. 2020;48(2):197–214. Available from: <https://dx.doi.org/10.1007/s12524-019-01069-4>.
- Mota GLA, Feitosa RQ, Coutinho HLC, Liedtke CE, Müller S, Pakzad K, et al. Multitemporal fuzzy classification model based on class transition possibilities. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2007;62(3):186–200. doi:10.1016/j.isprsjprs.2007.04.001.
- Berhane T, Lane C, Wu Q, Anenkhonov O, Chepinoga V, Autrey B, et al. Comparing Pixel- and Object-Based Approaches in Effectively Classifying Wetland-Dominated Landscapes. *Remote Sensing*. 2017;10(1):46–46. doi:10.3390/rs10010046.
- Clevers J, Kooistra L, van den Brande M. Using Sentinel-2 Data for Retrieving LAI and Leaf and Canopy Chlorophyll Content of a Potato Crop. *Remote Sensing*. 2017;9(5):405–405. Available from: <https://doi.org/10.3390/rs9050405>.
- Basu S, Ganguly S, Mukhopadhyay S, Dibiano R, Karki M, Nemani R. DeepSat: a learning framework for satellite imagery. In: and others, editor. *Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems*. 2015;p. 1–10. doi:https://doi.org/10.1145/2820783.2820816.
- Ienco D, Gaetano R, Dupaquier C, Maurel P. Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks. *IEEE Geoscience and Remote Sensing Letters*. 2017;14(10):1685–1689. doi:10.1109/LGRS.2017.2728698.
- Stromann O, Nascetti A, Yousif O, Ban Y. Dimensionality Reduction and Feature Selection for Object-Based Land Cover Classification based on Sentinel-1 and Sentinel-2 Time Series Using Google Earth Engine. *Remote Sensing*. 2020;12(1):76–76. Available from: <https://doi.org/10.3390/rs12010076>.
- Lopes M, Frison PL, Crowson M, Warren-Thomas E, Hariyadi B, Kartika WD, et al. Improving the accuracy of land cover classification in cloud persistent areas using optical and radar satellite image time series. *Methods in Ecology and Evolution*. 2020;11(4):532–541. Available from: <https://doi.org/10.1111/2041-210X.13359>.
- Ghassemian H. A review of remote sensing image fusion methods. 2016. Available from: <https://doi.org/10.1016/j.inffus.2016.03.003>.
- Otukei JR, Blaschke T, Collins M. Fusion of TerraSAR-x and Landsat ETM+ data for protected area mapping in Uganda. *International Journal of Applied Earth Observation and Geoinformation*. 2015;38:99–104. Available from: <https://doi.org/10.1016/j.jag.2014.12.012>.
- Noi PT, Kappas M. Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors*. 2018;18(1):18–18. Available from: <https://doi.org/10.3390/s18010018>.
- D P, J M. Developments in Landsat Land Cover Classification Methods: A Review. *Remote Sensing*. 2017;9(9):967–967. Available from: <https://doi.org/10.3390/rs9090967>.