

## Integrated mixture model and ensemble learning geographic object-based image analysis for road network extraction

Elaveni Palanivel & Shirley Selvan

To cite this article: Elaveni Palanivel & Shirley Selvan (2023): Integrated mixture model and ensemble learning geographic object-based image analysis for road network extraction, Journal of Spatial Science, DOI: [10.1080/14498596.2023.2217787](https://doi.org/10.1080/14498596.2023.2217787)

To link to this article: <https://doi.org/10.1080/14498596.2023.2217787>



Published online: 01 Jun 2023.



Submit your article to this journal 



Article views: 39



View related articles 



View Crossmark data 



# Integrated mixture model and ensemble learning geographic object-based image analysis for road network extraction

Elaveni Palanivel  and Shirley Selvan

Department of ECE, St. Joseph's College of Engineering, Chennai, India

## ABSTRACT

Delineation of road networks is often hindered by shadows, occlusions and spectrally similar objects. A hybrid Geographic object-based Image analysis (GeOBIA) technique that combines Mixture Model segmentation with Ensemble Learning algorithms to extract road networks is proposed. The Multispectral Local Dirichlet Mixture Model (MLDMM) highlights the built-up area. The bagging or subspace integrated ensemble learning-based classification makes the framework immune to overfitting and a novel statistical feature selection phase boosts the performance by 18%. Post-processing by path morphology achieves complete networks. MLDMM-SDEC proffers a precision of 99.86%, a recall of 90.90%, an F1-score of 95.23%, a detection quality of 94.73%, and an accuracy of 96.77%.

## ARTICLE HISTORY

Received 13 February 2023

Accepted 17 May 2023

## KEYWORDS

Mixture model;  
segmentation; road  
extraction; classification;  
GeOBIA

## Introduction

Road networks are essential infrastructure for connecting different parts of the world. Automatic extraction of road networks is crucial for several applications like urban design, vehicle navigation (Li *et al.* 2020), disaster response (Gupta *et al.* 2021), geospatial data integration (Kearney *et al.* 2020), etc., to name a few. The advent of contemporary and progressive remote sensing technologies has made high-resolution remote sensing imagery accessible for such applications. However, the presence of noise, occlusions caused by ambient structures (shadows, vehicles, vegetation, and trees), and road analogous features (car parking and railways) deter the extraction of accurate road networks.

Most of the conventional algorithms used for road network extraction, such as mathematical morphology (Zhu *et al.* 2005), mean shift segmentation (Miao *et al.* 2014), texture progressive analysis (Mena and Malpica 2005) and support vector machines (Lian *et al.* 2020), are pixel-centric to date. These algorithms utilize either pixel-dependent spectral information or the spatial variation of neighboring pixels within the same class. The texture that can provide unique identifiers is ignored. The accuracy of the results depends upon several initialization factors, such as seed points chosen by the user. Thus, these algorithms are limited in their performance and tend to generate more false positives.

Several of the earlier deep learning architectures that were proposed for road network extraction (Abdollahi *et al.* 2020) fell short in integrating crucial dependencies that highlight interclass and intraclass differences. Yang *et al.* (2022) developed an SDUnet for road network extraction. This network combined unique road features at multiple levels with prior global information. RADANET, proposed by Dai *et al.* (2023), utilized a pre-trained ResNet50 network integrated with a road augmentation module (RAM) and deformable attention module (DAM) to integrate geometric features with spectral data. Abdollahi *et al.* (2022) proposed an SC-RoadDeepNet that used an RRCNN with boundary learning to achieve the same. To perpetuate the connectivity of the extracted road network, a new measure, CP\_clDice, is proposed. All these networks integrated spatial, geometric, and spectral information to improve the accuracy of the road network extracted. These networks also performed well in the presence of shadows and occlusion. However, all these networks are pixel-centric, requiring a need for a large dataset for training. In any image, the percentage of road pixels is very low compared to the pixels belonging to other classes. This made the networks highly computation-intensive and hardware-dependent. Also, texture properties and their ability to differentiate between spectral and geometrically similar classes have not been explored.

An effective approach is the GeOBIA which considers objects rather than pixels as the basic unit of analysis (Chen *et al.* 2018). Thus, both spatial, spectral, and textural information represented by the object of interest is exploited to generate results. Here, the first phase is segmentation, where the pixels are essentially converted into objects. The features of each of the identified objects are extracted. Spatial, spectral, and textural characteristics suitable for the objects of interest are utilised. In the final phase, objects are classified as roads or non-roads. The major advantages (Blaschke *et al.* 2014) of this technique are

- The Segmentation step transforms the fundamental unit from a pixel to an image object. This makes the system more aligned with human cognition.
- The use of Image objects instead of pixels for analysis, helps reduce the ‘salt and pepper effect’ of the classified results as the internal homogeneity of complex classes is preserved.
- Also, high-level object features are extracted making any of the proposed algorithms more conducive to road extraction

The contribution of the paper is as follows

- A MLDMM algorithm for segmentation is proposed. This algorithm exploits the essential spectral and spatial information to achieve optimum segmentation and is robust to noise. The Dirichlet mixture model considers the number of clusters as a model parameter which is estimated automatically. So, there is no need for optimum parameter tuning.
- The normalisation of the pixel values removes the impact of differences in lighting. Thus, MLDMM also removes shadows and occlusions.
- The features for training the classification model are chosen statistically. This ensures that only the relevant features are taken into consideration thus eliminating any

redundancy in processing. This makes the framework less intensive in terms of computation and time.

- An ensemble of bagged and subspace discriminant classifiers with high bias-variance and less possibility of overfitting is used in the classification step. Ensemble learning is chosen as it is more robust, interpretable, flexible, less computationally expensive, and better at dealing with missing data and outliers.

The next section presents the literature survey followed by the workflow with an overview of the proposed framework. In the following section, segmentation, and classification algorithms are described. In the results and discussion section, the performance of the proposed algorithm with suitable results is discussed. The final section confers the conclusion.

## Literature survey

Road extraction approaches developed to date can be divided into different groups based on the categorisation criterion. The classification can be conventional or automatic based on the percentage of human involvement. It can also be supervised or unsupervised based on the need for a pre-defined dataset (Lian *et al.* 2020; Chen *et al.* 2022). Automated road acquisition is the preferred choice in the former category because of its surging need in several applications like urban planning, traffic management, and autonomous driving. In the latter category, the choice is not obvious. Supervised road extraction algorithms comprise those techniques that involve machine learning algorithms or deep learning networks for road extraction.

Miao *et al.* (2015) set forth one of the initial frameworks for road network extraction. Initial segmentation using the iterative self-organising data (ISODTA) analysis technique was followed by object-based Frangi's filtering (OFF) and shape filtering (OSF) for highlighting linear features. Finally, classification is achieved using Support Vector Machines (SVM). This framework did not exploit the entire structural and spectral information available. Also, the impact of shadow and occlusions were not discussed. Bakhtiari *et al.* (2017) proposed a semi-automatic algorithm incorporating SVM for classification and a full lambda schedule method for segmentation. Morphological operators were used in post-processing. The frameworks using machine learning algorithms proposed to date are not immune to noise and suffer from loss of performance because of occlusion. Guo and Wang (2020) proposed a self-supervised learning framework that introduced a positive sample selection method that combined spectral and shape features to retrieve road samples. A one-class Random Forest (RF) classifier is used to extract the road network. Though the level of automation is high in this algorithm, the accuracy of results is less when compared to deep learning techniques.

Maboudi *et al.* (2017) suggested an ant colony optimisation algorithm modified to work with structural and spectral information of objects for road network extraction. Yadav *et al.* (2020) put forward an OBIA fused fuzzy classification scheme. A membership function incorporating NDVI, Water Vegetation Index, and Intensity is proposed for road extraction. Both these algorithms could work only on multispectral data and not with the commonly available VHR images as the frameworks are dependent on the IR band. Also,

the algorithms perform poorly in the presence of shadows. Maboudi *et al.* (2018) suggested a hybrid algorithm merging a fuzzy inference system with ant colony optimisation. A road model with all the structural, spatial, and textural descriptors represented by the associated membership functions is developed. However, the fuzzy rule set proposed could not be generalised for all input images and the membership function also needs to be adjusted for different datasets.

Several deep learning-based networks such as the deep window technique (Lian *et al.* 2020), a multi-scale and multi-task deep learning network (Lu *et al.* 2019), recurrent convolution neural network U-net (RCNN-UNET) (Yang *et al.* 2019) and the deep residual U network (Zhang *et al.* 2018) have been proposed for road extraction that achieves varying degrees of excellence in terms of performance (Abdollahi *et al.* 2020) (Abdollahi et al. 2018). Yang *et al.* (2022) proposed a new CNN-integrated TransRoadnet network for road extraction, overcoming the inability of initial CNN networks to capture contextual information. This was achieved by introducing a swin transformer. The majority of the deep learning networks proposed are pixel centric. For the first time, a multiscale object-based convolutional neural network (multi-OCNN) was proposed by Martins *et al.* (2020). Initial object identification was carried out by the mean shift algorithm. This was followed by training 6 neural networks each processing data in different patch sizes. These networks were then used to classify the input image into 10 different classes including roads, buildings, and shadows. Here the concept of objects is used to reduce the number of predictions only. Also, the network doesn't perform well in distinguishing between roads, buildings, and sand. Despite several favourable developments in recent years, supervised techniques still have a lot of scope for improvement. Deep learning methods process images in the form of rectangular shape patches only, due to which object boundaries extracted may not be continuous. Despite several merits, all deep learning algorithms suffer from the need for a large pre-labelled dataset which is time-consuming, costly, and tedious. Also a large data storage capacity and costly high configuration hardware is also needed.

Unsupervised road extraction techniques overcome some of the obstacles faced by supervised learning, namely the need for a large dataset for training and computational complexity respectively. Liu *et al.* (2019) developed a semi-supervised high-level feature selection framework for road centerline extraction. A feature learning framework with MRF is proposed. Gabor filters and non-maxima suppression are integrated with the ridge transversal method to extract the road centerlines. Li *et al.* (2018) utilised the gaussian mixture model to extract roads. The algorithm comprises of superpixel segmentation, feature description, homogeneous region merging, clustering via the Gaussian mixture model, and outlier filtering. This method is fully automatic. Despite all the improvements, the performance of unsupervised algorithms is less than its supervised counterparts.

The majority of the Segmentation algorithms developed to date are either spectral or spatial-based. Kotaridis and Lazaridou (2021) True to their classification, the traditional algorithms utilise either the spectral or the spatial data only to achieve segmentation. The spectral-based algorithms are all pixel centric and hence are out of our scope. On the other hand, spatial algorithms such as edge and region-based algorithms work with the spatial information only and do not exploit the spectral or textural information. Deep learning-based segmentation techniques developed to date still struggle to combine all the high-level semantic

information with the low-level spatial details. The intimidating size of the imagery compared with the meagre ground truth information available is also a deterrent. Also, the huge number of parameters that need to be tuned makes it an impractical pre-processing step (Minaee *et al.* 2021). Since segmentation is key to the success of any GEOBIA framework, a competent algorithm that exploits all available information to dispense accurate objects is imminent.

Model-based segmentation algorithms are a robust choice as they can utilise statistical prior knowledge along with structural, spectral, and textural information to achieve optimum segmentation. Song *et al.* (2017) used a Dirichlet process mixture model (DPMM) with an MRF framework for the segmentation of PolSAR images. The DPMM makes the system immune to noise. Mantripragada and Qureshi (2022) introduced a Dirichlet process gaussian mixture model-based segmentation technique for noisy hyperspectral images. They compared their proposed algorithm with several state-of-the-art segmentation techniques and have proved experimentally that mixture model-based segmentation techniques provide better results even in the presence of noise.

Attempts have been made to combine the best features of both supervised and unsupervised algorithms. Chen *et al.* (2019) put forth a combination of the Local Dirichlet Mixture Model (LDMM) with ResNet for road extraction. This algorithm achieved good results with reduced false positives. Thus, coupling an unsupervised segmentation algorithm with a supervised classification algorithm can achieve a good compromise.

This paper builds on this idea and integrates the merits of mixture model-based segmentation into the GEOBIA framework. A novel unsupervised multispectral local Dirichlet mixture model (MLDMM) is integrated with subspace discriminant ensemble learning (SD-EL).

### **Workflow of multispectral local Dirichlet mixture model using subspace discriminant ensemble classifier**

The workflow as shown in Figure 1, consists of two broad phases namely GeOBIA followed by connected component analysis for acquiring continuous road networks.

In GeOBIA, initially, the input image is segmented. MLDMM segments the image into built-up and non-built-up areas. In this step, vegetation interspersed within and around the road network along with waterbodies is eliminated. As a result, the connected components representing the built-up area in the image are highlighted. The next step is to identify those connected components that comprise the road networks. To achieve this, the features of the objects are extracted in the second stage. Spectral, geometric, and textural features unique to roads are adopted. These features are then used to train a plethora of ML algorithms and the results are compared. Ensemble learning algorithms are identified as the best fit. The trained algorithm is then used to classify the input image components into roads and non-roads.

The road network obtained at the GeOBIA stage's end is discontinuous. The gaps are filled and the final output is updated using connected component analysis, which employs path Morphology to achieve this. Thus, the final road network is extracted.

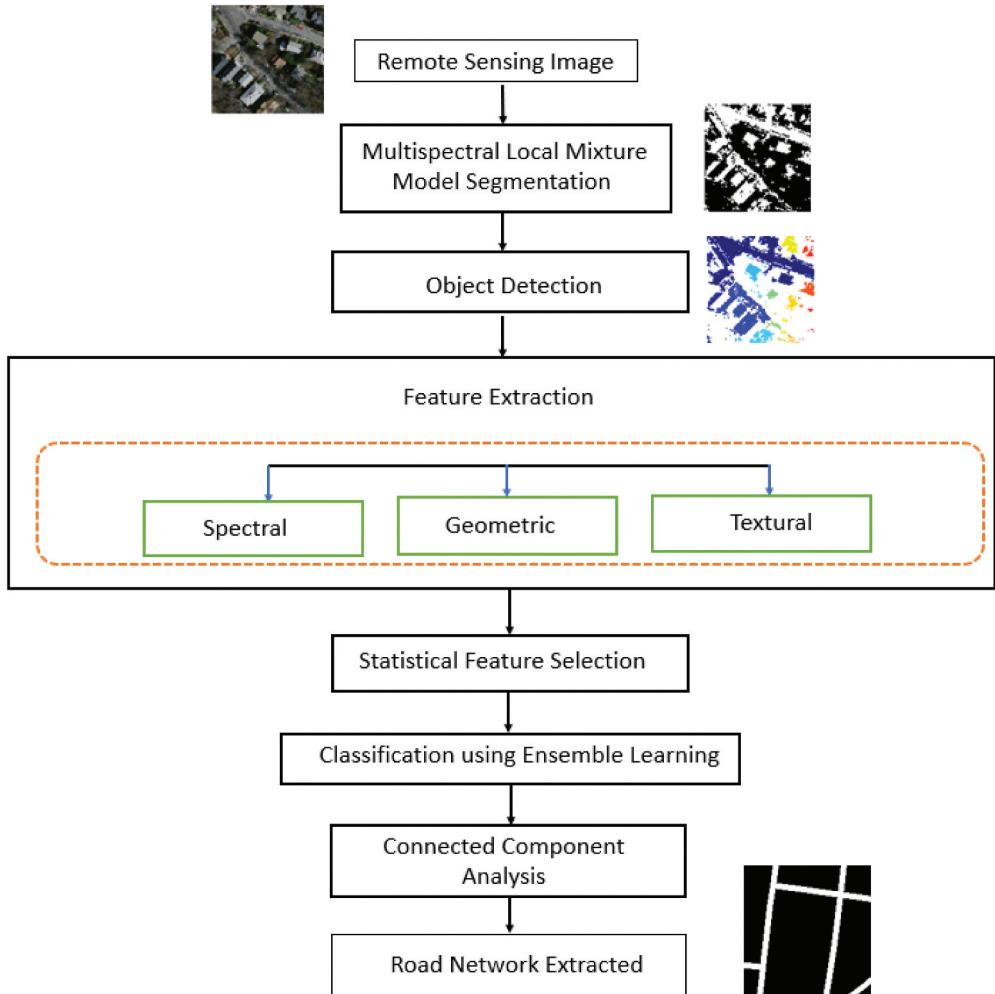


Figure 1. Workflow of the proposed algorithm.

### Segmentation

Segmentation associates a group of neighbouring pixels based on some preconceived parameters. As a result, moderately homogenous regions are highlighted which are the objects of interest. However, the accuracy of all segmentation algorithms relies on three crucial criteria: congruence within a segment, the contradiction between adjacent regions, and similarity in shape (Kotaridis and Lazaridou 2021). These factors can be appraised efficiently only by merging both the spatial and spectral information. Mixture model-based segmentation algorithms easily achieve this. These integrate spatial smooth data with spectral information and achieve better segmentation (Hu *et al.* 2018).

The MLDMM segmentation method is adopted in the proposed algorithm. The original algorithm proposed by Chen *et al.* (2019) is modified to utilise all the spectral data available rather than just the grey value information. As a result, the entire spectral information in combination with the spatial data is exploited. The input image is divided

into patches and MLDDMM is employed on each of the patches individually. The output of this stage highlights the most likely road segments.

### ***Dirichlet mixture model with spatial constraints***

In MLDDMM, the R, G, and B values of a pixel are considered. Thus, every input pixel  $P_i$  represents three-dimensional data  $(p_r, p_g, p_b)$ . In this algorithm, a pixel is modelled as a three-dimensional vector whose architect is a Dirichlet distribution with parameters  $\varpi = (\varpi_r, \varpi_g, \varpi_b)$  represented in Equation (1)

$$\text{Dir}\left(\frac{\rightarrow}{P} \mid \frac{\rightarrow}{\varpi}\right) = \frac{\Gamma(\sum_{i=1}^3 \varpi_i)}{\prod_{i=1}^3 \Gamma(\varpi_i)} \prod_{i=1}^3 p_i^{\varpi_i-1} \quad (1)$$

According to the finite Dirichlet mixture model, a pixel is an amalgamation of several parent Dirichlet distributions with varying degrees of contribution. The proportion of the contribution of each distribution is represented by a mixing co-efficient  $a = (a_1, a_2, \dots, a_J)$  where  $J$  represents the number of parent Dirichlet distributions of a pixel. When the information in the spatial correlation of adjacent pixels in an image is also considered, the proportion parameter of the first pixel is modified as  $a_1 = (a_{11}, a_{12}, \dots, a_{1J})$ . This information is important as the probability of adjacent pixels belonging to the same cluster is high (Blekas *et al.* 2005). If the total number of pixels in an image is  $M$ , they can be represented as in Equation (2)

$$p(X|\varpi, a) = \prod_{m=1}^M \left[ \sum_{j=1}^J a_{mj} \text{Dir}(P_m|\varpi_j) \right] \quad (2)$$

The variable indicating the class to which the pixel belongs is represented by  $\varphi$ . Characteristics neighbourhood information is introduced by replacing the value of a pixel with the average of its neighbours as in Equation (3)

$$\overline{\varphi}_{mj} = \frac{\sum_{i \in \Theta} \varphi_{ij}^{(u-1)}}{|\Theta_m|} \quad (3)$$

where  $\Theta$  denotes the neighborhood of the pixel. A  $3 \times 3$  window is chosen the same as the original algorithm. Variational learning is used to determine  $a$  and  $\varphi$  as discussed in Chen *et al.* (2019). Based on the value of  $a$  and  $\varphi$ , the number of clusters and the class of each of the pixels are determined. Thus, the patches representing the built-up area are extracted. Each of the independent patches is considered an object for further processing.

### ***Road feature extraction***

Humans perceive roads by utilising their geometric (road shapes), radiometric (colour or grey levels), topological (road intersections), and textural (regional) characteristics. However, roads have a unique spectral signature in comparison with the background (Das *et al.* 2011). Thus, it is ideal to model a framework, by integrating human-centric and spectral features. Mean and standard deviation are calculated to represent the spectral data. After identifying the objects, the mean and standard deviation of the (R, G, B) value of the pixels in the objects are calculated. The pixel values are normalised to remove the distraction of lighting variations ensuring that the underlying feature is effectively captured by the classification algorithm (Pan *et al.* 2018). These parameters are calculated from all the frames to benefit from the complete spectral information.

The other salient feature of roads is their geometric properties. Roads are essentially elongated objects with inherent local linearity. The Area, Length, Width, and number of pixels are indicators of geometry. Along with these, the Complexity Rate ( $\alpha$ ), Aspect Ratio ( $\beta$ ) of the bounding rectangle, and Extent ( $\gamma$ ) (Soni *et al.* 2020) are also used to better interpret the geometry of roads.

Texture is a descriptor that characterises coarseness, contrast, regularity, and directionality in the spatial distribution of an image. Despite significant ambiguity in its definition to date, the examination of texture augments the performance of road extraction in remote sensing images (Chaki and Dey 2020). Our framework selects second-order statistical features for texture extraction using the Grey-Level Co-occurrence Matrix (GLCM). GLCM is invariant to monotonic grey-level transformation and orientation, making it the best fit for road network extraction (Ramola *et al.* 2020). However, it is computationally intensive since it is calculated at the pixel level. This snag is beaten in our framework as we calculate the GLCM features only for the pixels of the objects extracted in the segmentation phase.

The features used in the proposed algorithm are listed in Table 1 below. These features have been compiled based on a detailed survey of the earlier literature and taking into account the innate properties of a typical road network.

The features are then analysed statistically to determine the significant features. This helps reduce the complexity of the proposed algorithm by reducing the amount of data that needs to be processed. Also, this contributes to improved accuracy of the output. The p-value of the features is calculated and only the most significant of them are chosen as inputs for further classification. Initially, a population of 1000 patches is compiled. All the listed features are extracted from the patches to create a database. Both T and Z tests are implemented. The T-tail test is calculated for a small sample size and the Z-test is for a larger sample size. The p-value of the features is listed in Table 2. 25 samples are used to carry out the T-test and all 1000 samples are used to carry out the Z-test. When the p-value is less than 0.05, the null hypothesis is rejected in T-test. In the case of the z two-tailed test, the z value has to be greater than 1.96 or less than -1.96 for the null hypothesis to be rejected.

**Table 1.** Features extracted for road network extraction.

S. No	Category		Features
1	Spectral Features	a	Mean $\mu = \sum_{i=1}^N x_i / N$
		b	Standard Deviation $\sigma = \sqrt{\sum (x - \mu)^2 / n}$
2	Geometric Features	a	Complexity Rate ( $\alpha$ ) $\alpha = \text{Area} / (\text{Perimeter})^2$
		b	Aspect Ratio ( $\beta$ ) of the bounding rectangle $\beta = \frac{\text{Length of BR}}{\text{Width of BR}}$
		c	Extent ( $\gamma$ ) $\gamma = \frac{\text{Area of Object}}{\text{Area of the min BR}}$
3	Texture Features	a	Contrast
		b	Energy
		c	Dissimilarity
		d	Homogeneity
		e	Correlation
			Features extracted from the GLCM Matrix

**Table 2.** Results of statistical analysis.

S. No	Category		Features	P-Value (t-tail test)	Z-value (z two-tailed test)
1	Spectral Features	a	Mean/Area	0.011153	6.119304
		b	Standard Deviation	0.02546	6.23546
2	Geometric Features	a	Major Axis	0.026882	6.977911
		b	Minor Axis	0.041108	6.100544
		c	Convex Area	0.018667	7.028392
		d	Complexity Rate ( $\alpha$ )	0.056915	0.665905
		e	Aspect Ratio ( $\beta$ ) of the bounding rectangle	0.509793	-0.13895
		f	Extent ( $\gamma$ )	0.065116	0.991175
		g	Area of Bounding rectangle	0.682488	-5.059856
3	Texture Features	a	Contrast	0.031965	6.248826
		b	Energy	0.604271	-7.24661
		c	Dissimilarity	0.065432	5.43213
		d	Homogeneity	0.055701	-6.24825
		e	Correlation	0.61661	-1.49071

All the spectral features are significant. In the case of geometric features, the aspect ratio and area of the bounding rectangle are proved insignificant. This is because, the road network is a long feature and a rectangle, bounding the entire road network does not represent useful information. Because of the same phenomenon, the area of the bounding rectangle is also insignificant. All texture features are significant except correlation and energy.

Correlation is a measure of the joint probability occurrence of specified pixel pairs. The correlation of texture indicates repetitive variation occurring in an object. Roads, on the other hand, predominantly have a gradual flowing texture; hence, correlation does not hold any significant information in this case. Energy also known as uniformity or angular second moment quantifies the global homogeneity or uniformity of texture. Rather, more importance should be vested in measuring the local variations in texture since it would contribute more towards identifying the boundaries of road networks. Thus, Energy is omitted while contrast is statistically proven as a significant feature.

## Classification

The goal of classification is to assess a sensed pattern and assign it to the most suitable class (Duda and Hart 2006). The significant features are extracted from the objects obtained at the end of MLDDM segmentation. Literature suggests several machine learning (ML) algorithms that can differentiate pixels into road and non-road categories, (Kanevski *et al.* 2009; Kroese *et al.* 2019) but their fitness to manipulate remote sensing data is the decisive criterion. All the ML algorithms were surveyed and Support vector machines (SVMs), Decision trees (DT) and Ensemble learning (EL) were chosen to validate the performance of the proposed algorithm.

Ensemble learning involves combining multiple weak classifiers based on a defined set of rules thus improving the overall performance (Zhou 2012). Rather than choosing a single algorithm and trying to fit the model to our data, ensemble learning provides the flexibility to utilise the strength of multiple classifiers. The commonly used ensemble learning techniques are bagging, boosting, and stacking. The bagging technique is preferred for our application as multiple models can be trained independently and

parallelly. This reduces the processing time. Thus, bagging is the preferred algorithm. But instead of the most commonly used Random Forest bagging classifier (RFBC) (Zhang *et al.* 2022), the subspace discriminant classifier (SDC) is used as it is designed to maximise the difference between the means of the classes in the selected feature subspace, which can lead to better discrimination between classes compared to RFBC, which does not explicitly optimise for class separation. SDC also has a low variance which helps avoid overfitting.

After classification, path closings and path openings are used to fill the gaps in road networks (Valero *et al.* 2010). These advanced morphological operators are highly tractable and hence can fill the structural information represented by both orthogonal and slightly arched features. This enables the accurate extraction of road networks without the loss of crucial curvilinear information.

## Results and discussion

The proposed algorithm is verified experimentally using the Massachusetts Road (MR) Dataset. This dataset consists of 1171 aerial images of the state of Massachusetts. Each image covers an area of 2.25 square kilometres. The size is  $1500 \times 1500$  pixels. All images are rescaled to a resolution of 1 pixel per square metre (Mnih 2013). All images have corresponding ground truth associated with them. Patches from 1107 images are used for training and 13 images are used for testing. Test images from the dataset along with its ground truth are shown in Figure 2

Initially, image patches of size  $100 \times 100$  are extracted from training images. Patches, where multiple classes overlap are chosen to show the robustness of the proposed algorithm in extracting road networks even in the presence of shadow and occlusion. The results of 4 patches are discussed in detail. Patch 1 consists of roads interspersed with buildings and trees. The buildings also cast shadows. Patch 2 shows a road junction occluded by trees with buildings nearby. Patch 3 encloses a portion of a road that crosses a water body. Finally, patch 4 consists of roads with varying widths occluded by vegetation and surrounded by buildings.

MLDMM segmentation is applied to the patches. The indirect indicator variable  $\varphi$  is randomly initialised to either 0 or 1. The factor indicating the proportion of the contribution of the parent distributions  $\alpha$  is also initialised randomly to any value between 0 and 1



**Figure 2.** Test data from MR dataset (a) Site A (b) Ground truth of site A (c) Site B (d) Ground truth of site B.

in the simulation. In MLDDMM, the final optimum value of the parameters is independent of the initialised value. Also, the algorithm is robust to noise and shadows because of the spatial constraints imposed (Hu *et al.* 2018). The results are shown in Figure 3. Column (a) consists of all 4 patches. Column (b) is the segmented output. Column (c) highlights all the connected components in the output, identified as objects. These objects highlight the region of interest.

The segmentation results displayed in column (b) of Figure 3 demarcate the built-up and non-built-up areas. In all the patches, vegetation is visibly filtered out. The shadows in Patch 1 are filtered out and only the building and the road regions are highlighted as objects. The river in patch 3 is also eliminated. From the output, it is clear that MLDDMM effectively removes the occlusions and shadows from the road structure. From patch 1 and patch 4 output, it can be deduced that buildings and road structures are both highlighted at the end of segmentation. To differentiate buildings from roads, geometric and texture properties are utilised. When the perimeter and Major axis are used, buildings get eliminated. This is represented by outputs in columns 4 and 5. Thus, the objects of interest are extracted at the end of the segmentation phase.

Segmentation is followed by feature extraction. Features that have been deemed necessary are extracted. 6000 objects extracted from 1250 image patches are used.

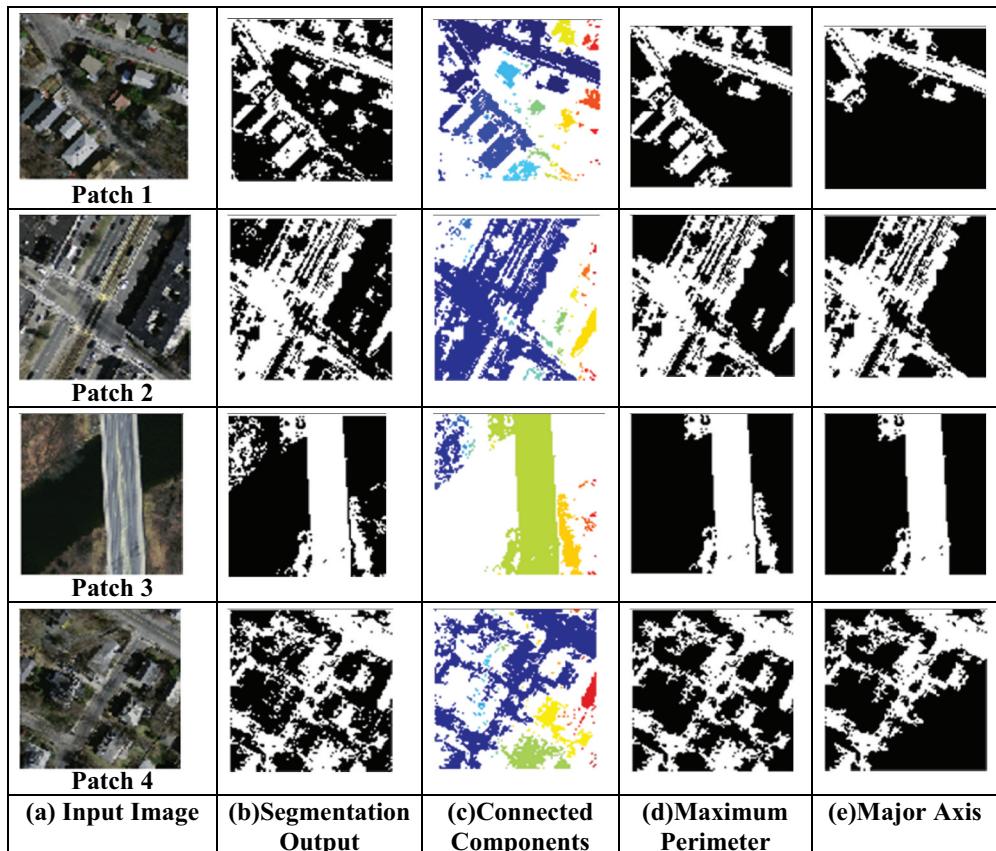


Figure 3. Segmentation results of the patches of interest.

After extracting the features of all objects, a statistical Z-test and two-tailed T-test are used to identify statistically significant features. All other features are eliminated. The chosen statistically significant features are then used to train the ML algorithms. Of the 6000 objects, 2300 are identified as road objects and the rest as non-road objects. The results of the statistical analysis are shown in [Table 2](#). The p values and z values for all the features are listed. The significant features are highlighted.

Feature selection is followed by classification. An Ensemble of machine learning models is employed for classification. Bagged decision trees, Subspace discriminant, and subspace KNN are the chosen models. The ML algorithms used are listed in [Table 3](#). These algorithms attempt to classify objects as roads or non-roads. The subspace dimension refers to the subset of features used in each subspace of the classifier. If the subspace dimension is very high, it may result in an unnecessary increase in the complexity of the model and overfitting, while a very small subspace dimension may result in underfitting. Hence a median value of 5 is chosen. After extensive experimentation, an optimum value of 30 learners with a subspace dimension of 5 and a learning rate of 0.1 is chosen for the ensemble learners. In the decision tree bagged ensemble learning (BTEC) a maximum of 41 splits and in subspace KNN learning, a maximum of 20 splits is the optimum value.

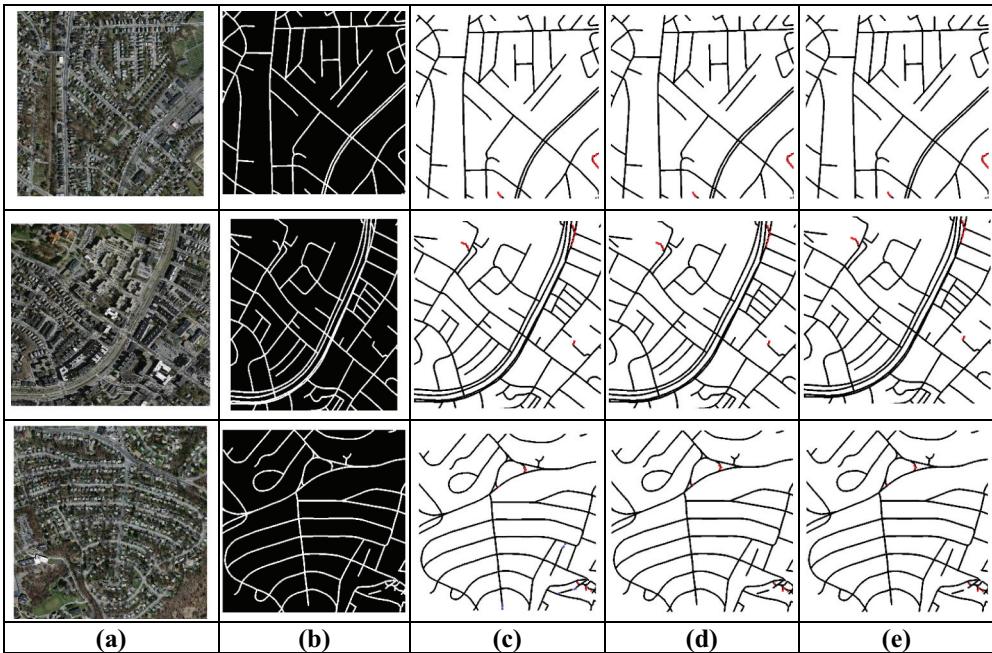
The extracted road networks are shown in [Figure 4](#). [Figure 4\(b\)](#) shows the ground truth while [Figure 4\(c–e\)](#) are extracted networks of MLDMM-SDEC, MLDMM-KNN-EC, and MLDMM-BTEC respectively. The correctly identified road objects are considered true positive (TP). When the actual road objects are classified as non-road, they fall under false negatives (FN). When correctly detected as non-road objects, they become the true negatives (TN), whereas the incorrectly identified road objects become false positives (FP). In [Figure 4](#), TP, FP, and FN are represented by black, blue, and red colours respectively. The performance of the ML algorithms is assessed using the following metrics.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

**Table 3.** Average performance metrics of the proposed algorithm.

S. No	Algorithm	With Feature Selection			
		Precision	recall	F1 Score	Detection Quality
1	Coarse Tree	0.812	0.893	0.851	0.765
2	Linear Discriminant Analysis (LDA)	1	0.928	0.963	0.795
3	Quadratic SVM	1	0.838	0.9118	0.865
4	Cubic SVM	0.968	0.5	0.659	0.789
5	Medium Gaussian SVM (MG-SVM)	0.967	0.871	0.917	0.811
6	Coarse Gaussian SVM (CG-SVM)	0.968	0.833	0.895	0.75
7	Bagged Trees Ensemble Classifier (Bagged Tree EC)	0.967	0.865	0.913	0.882
8	Subspace Discriminant Ensemble Classifier (SDEC)	<b>0.998</b>	0.909	0.952	<b>0.947</b>
9	Subspace KNN Ensemble Classifier (KNN-EC)	0.967	<b>0.947</b>	<b>0.957</b>	0.906



**Figure 4.** (a) Input image (b) Ground truth (c) Output of MLDMM-SDEC (d) Output of MLDMM-KNN-EC (e) Output of MLDMM-BTEC.

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

$$Detection\ Quality = \frac{TP}{TP + FP + FN} \quad (7)$$

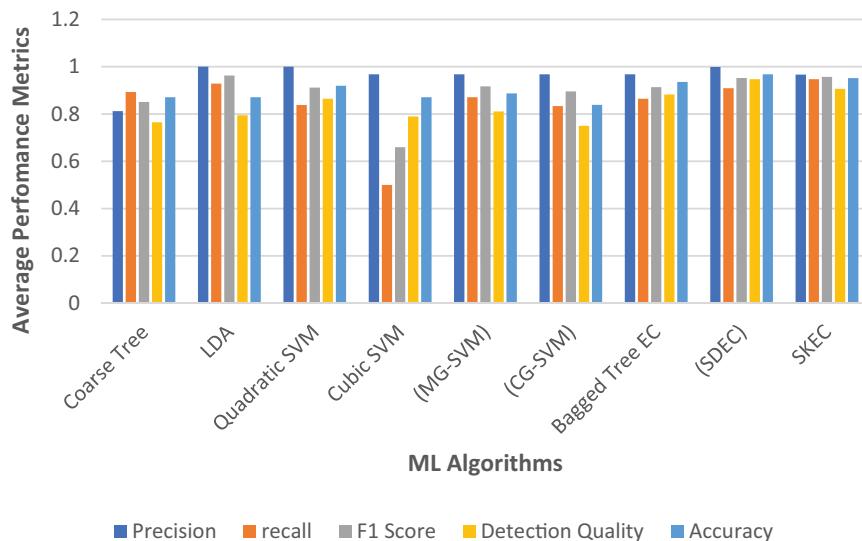
$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

Recall is a measure of the completeness of the road network extracted. Precision or correctness measures the accuracy of the output. F1 score is the harmonic mean of precision and recall. The performance metrics of the algorithms producing the best results are shown below in [Table 3](#).

[Figure 5](#) compares the performance metrics of all the proposed algorithms. Ensemble learning integrated algorithms especially, subspace discriminant ensemble learning displays the best metrics when compared to other algorithms. This shows the superiority of combining multiple learning algorithms rather than trying to fit a single model.

### Ablation study 1

An ablation study is carried out to indicate the importance of feature selection. The same classification algorithms are trained using all available features rather than just the statistically significant ones. The metrics obtained are listed in [Table 4](#).

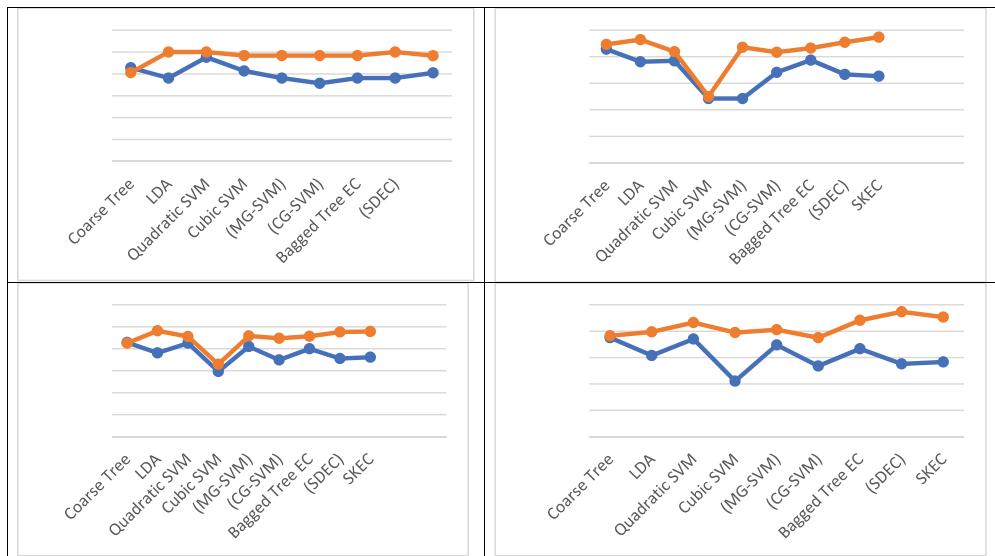


**Figure 5.** Comparison of performance metrics of the proposed algorithms.

**Figure 6** analyses the impact of feature selection on the performance metrics. The orange graph refers to the metrics with feature selection while the blue graph refers to the metrics without feature selection. It can be inferred that there is a marked improvement in the performance metrics after feature selection. On average, precision has increased by 17.83%, recall by 12.05%, F1 score by 14.42%, detection quality by 25.32%, and accuracy by 20.29%. This is because the feature selection helps identify only the relevant features. This reduces the dimension of the data thereby reducing the computation complexity and time. On the other hand, since only the pertinent features are processed the possibility of error in the results diminishes remarkably as proved by the statistics. Ensemble learning algorithms perform better than the rest as they utilise multiple learners to classify the data. The subspace discriminant ensemble classifier performs best for the chosen Massachusetts dataset.

**Table 4.** Average performance metrics of machine learning algorithms without feature selection.

S. No	Algorithm	Without Feature Selection				
		Precision	recall	F1 Score	Detection Quality	Accuracy
1	Coarse Tree	0.857	0.857	0.857	0.75	0.762
2	Linear Discriminant Analysis (LDA)	0.762	0.762	0.762	0.615	0.564
3	Quadratic SVM	0.952	0.7691	0.851	0.741	0.793
4	Cubic SVM	0.827	0.485	0.592	0.421	0.738
5	Medium Gaussian SVM (MG-SVM)	0.762	0.889	0.820	0.696	0.690
6	Coarse Gaussian SVM (CG-SVM)	0.714	0.682	0.698	0.536	0.619
7	Bagged Trees Ensemble Classifier (Bagged Tree EC)	0.762	0.774	0.8	0.667	0.704
8	Subspace Discriminant Ensemble Classifier (SDEC)	0.762	0.667	0.711	0.552	0.690
9	Subspace KNN Ensemble Classifier	0.809	0.653	0.723	0.567	0.714



**Figure 6.** Impact of feature selection over average performance metrics of machine learning algorithms.

On comparing the performance metrics of the algorithms with feature selection, as shown in Figure 6, ensemble learning algorithms perform better than the rest as they utilise several learners to classify the data. The subspace discriminant ensemble classifier performs best for the chosen Massachusetts dataset.

### Ablation study 2

The second ablation study is carried out to analyse the impact of different feature categories. Initially, each category of features namely spectral, spatial, and geometric features are independently used to train the classifiers. Then different combinations of features namely spectral and geometric, geometric and textural, spectral and textural, and finally all three feature categories are used to train the classifiers. The performance metrics are listed in Table 5. The detection quality is very poor when only one category of features is used for training. Optimum results are obtained only when all the statistically significant features are involved.

Figure 7 shows the parallel coordinate plots of the ensemble learning algorithms when only textures were used for training. The blue lines represent the non-road objects and the orange line represents the road objects. From the plot, it is clear that homogeneity and energy do not provide a clear separation of road and non-road objects, while the other textural features show a clear demarcation. This further supports the decision that correlation and energy are not statistically significant while identifying road networks.

### Topological evaluation

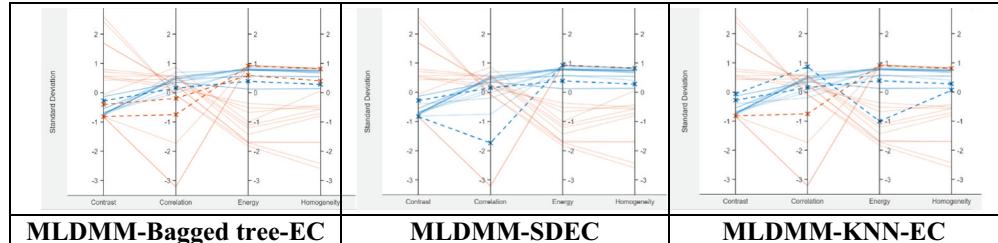
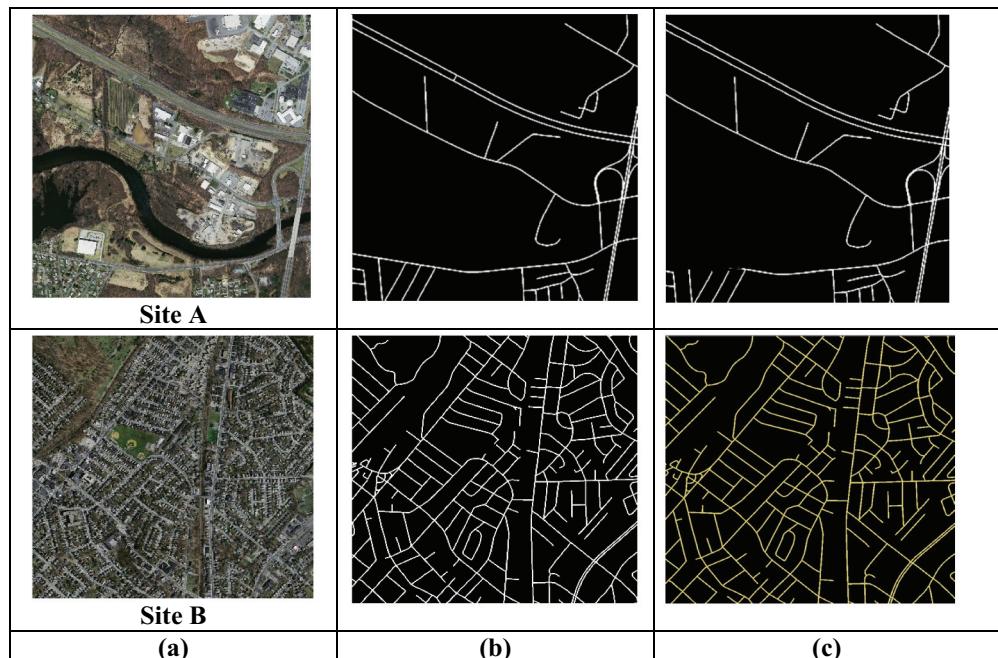
The efficacy of the road network extracted can be established by its topological evaluation (Wiedemann 2003). Topological completeness and topological correctness give

**Table 5.** Performance metrics of ensemble learning classifiers with different combinations of feature categories.

Feature Category	MLDMM-BT-EC			MLDMM-SDEC			MLDMM-KNN-EC					
	Precision	Recall	F1-score	Detection Quality	Precision	Recall	F1-score	Detection Quality	Precision	Recall	F1-score	Detection Quality
Spectral features only	0.812	0.867	0.839	0.722	0.818	0.6	0.692	0.529	0.75	0.8	0.774	0.631
Geometric Features only	0.928	0.867	0.896	0.812	0.9167	0.733	0.815	0.687	0.917	0.733	0.815	0.687
Textural Features only	0.706	0.8	0.75	0.6	0.667	0.8	0.727	0.571	0.706	0.8	0.75	0.6
Spectral+Geometric	0.812	0.867	0.839	0.722	0.917	0.733	0.815	0.687	0.875	0.933	0.903	0.823
Spectral+Textural	0.812	0.867	0.839	0.722	0.923	0.8	0.857	0.75	0.929	0.867	0.896	0.812
Geometric+spectral	0.875	<b>0.933</b>	<b>0.903</b>	0.823	0.867	0.867	0.765	0.765	0.8	0.8	0.8	0.667
All statistically significant features	0.968	0.865	0.913	0.882	<b>0.999</b>	<b>0.909</b>	<b>0.952</b>	<b>0.947</b>	<b>0.967</b>	<b>0.947</b>	<b>0.957</b>	<b>0.906</b>

**Table 6.** Topological evaluation metrics for site A and site B.

S. NO	Test data	$NP_{ground}$	$NP_{result}$ (MLDMM-SDEC)	$NP_{homologous}$	$NP_{extra}$	Topological Completeness	Topological Correctness
1	Site A	30	27	27	27	90%	100%
2	Site B	137	136	128	128	93%	94.11%

**Figure 7.** Parallel coordinate plots of classifiers when only texture features are used for training.**Figure 8.** (a) Input image (b) Ground truth (c) Output of MLDMM-SDEC.

a measure of the extent of connections, that the extracted road network has when compared to the ground truth. If the connections in the extracted road network are more than that in the ground truth, then it is topological incorrectness and if less, then it is topological incompleteness.

Topological Completeness is the ratio of the homologous node pairs ( $NP_{homologous}$ ) in the extracted result to the total number of node pairs ( $NP_{ground}$ ) in the ground truth. 100%

is the ideal result. If the value is less, it reflects the fragmentation in the results. Topological Completeness is given as per Equation (9).

$$\text{Topological Completeness} = \frac{NP_{homologous}}{NP_{ground}} \quad (9)$$

Topological correctness determines if there are node pairs in the extracted result that are not present in the ground truth. To calculate this metric, all node pairs in the extracted result ( $NP_{result}$ ) are calculated initially. Then, those node pairs in  $NP_{result}$  that matches with the ground truth ( $NP_{extra}$ ) are calculated. A ratio of these values gives the topological correctness. If the value is less than the maximum value of 100%, it indicates the presence of wrong connections in the extracted result.

The topological completeness and correctness for Test site A and B are tabulated in [Table 6](#). The ground truth and the road network extracted for sites A and B are displayed in [Figure 8](#).

$$\text{Topological Correctness} = \frac{NP_{extra}}{NP_{result}} \quad (10)$$

### **Comparison with existing techniques**

The Dirichlet mixture model-based ensemble learning algorithms proposed is compared with some recently proposed road extraction algorithms that use the same Massachusetts dataset for verifying their results.

Soni *et al.* (2020) combined a couple of shape features with LS-SVM for the initial segmentation of images into road and non-road objects. They fine-tuned their result by using conventional mathematical morphology. This method was not robust to occlusion and there was no discussion about the impact of shadows. Abdollahi *et al.* (2021) proposed an object-based framework for road network extraction. A multiresolution segmentation is followed by classification using decision trees, SVM and Knn. A standard set of features have been used for training these classifiers without any statistical or logical analysis. Also, there is no discussion regarding shadows or occlusions. In the case of deep learning algorithms, SC-RoadDeepNet proposed by Abdollahi *et al.* (2022) produced a robust road network. Since all this literature also used the Massachusetts dataset, the metrics are compared in [Table 7](#).

**Table 7.** Comparison with existing algorithms.

Algorithm	Precision	Recall	F1-score
MLDMM-BTEC	0.9677	0.8648	0.9134
MLDMM-SDEC	<b>0.998654</b>	0.878788	0.935484
MLDMM-KNN-EC	0.9667	<b>0.9473</b>	<b>0.9569</b>
Soni <i>et al.</i> (2020)	0.8354	0.8566	0.8897
Abdollahi <i>et al.</i> (2021)	0.8954	0.8765	0.8988
Abdollahi <i>et al.</i> (2022)	-	-	0.8933



## Conclusion

With the fast evolution of remote sensing technology and its rising importance in urban development and planning, the need for an optimum algorithm to extract road networks is vital. The MLDMM-SDEC algorithm proposed here fills this void. It is a modern GeOBIA technique that integrates the best features of unsupervised and supervised algorithms. The MLDMM segmentation algorithm carries out the dual role of removing the vegetation and water bodies in the background and removing shadows and occlusion. The built-up area is identified and from it, the objects of interest are extracted. Features are extracted only from these objects of interest. Using statistical analysis to select features for training the classifiers that extract the road network has been attempted for the first time to the knowledge of the author. This eliminates the need for a large dataset and reduces the computational complexity of the framework. The significance of feature selection and the choice of features has also been verified experimentally in the two ablation studies. Several machine learning algorithms are trained and tested with the features. Ensemble learning is chosen as it gives the best results. Performance, Topological completeness, and correctness of MLDMM-SDEC are evaluated and are found to be better than the existing conventional and machine learning algorithms but comparatively less than that of deep learning techniques. However, there is still scope for improvement. Though the segmentation phase is independent of any initialisation, some parameters are chosen experimentally in the classification stage. Automatic optimisation of these parameters can be an area for future analysis.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Elaveni Palanivel <http://orcid.org/0000-0002-5006-338X>

## References

- Abdollahi, A., et al., 2020. Deep learning approaches applied to remote sensing datasets for road extraction: a state-of-the-art review. *Remote Sensing*, 12 (9), 1444. doi:[10.3390/rs12091444](https://doi.org/10.3390/rs12091444)
- Abdollahi, A. and Pradhan, B., 2021. Integrated technique of segmentation and classification methods with connected components analysis for road extraction from orthophoto images. *Expert Systems with Applications*, 176, 114908. doi:[10.1016/j.eswa.2021.114908](https://doi.org/10.1016/j.eswa.2021.114908)
- Abdollahi, A., Pradhan, B., and Alamri, A., 2022. SC-roaddeepnet: a new shape and connectivity-preserving road extraction deep learning-based network from remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–15. doi:[10.1109/TGRS.2022.3143855](https://doi.org/10.1109/TGRS.2022.3143855)
- Bakhtiari, H.R.R., Abdollahi, A., and Rezaeian, H., 2017. Semi-automatic road extraction from digital images. *The Egyptian Journal of Remote Sensing and Space Science*, 20 (1), 117–123. doi:[10.1016/j.ejrs.2017.03.001](https://doi.org/10.1016/j.ejrs.2017.03.001)
- Blaschke, T., et al., 2014. Geographic object-based image analysis—towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 180–191. doi:[10.1016/j.isprsjprs.2013.09.014](https://doi.org/10.1016/j.isprsjprs.2013.09.014)
- Blekas, K., et al., 2005. A spatially constrained mixture model for image segmentation. *IEEE Transactions on Neural Networks*, 16 (2), 494–498. doi:[10.1109/TNN.2004.841773](https://doi.org/10.1109/TNN.2004.841773)

- Chaki, J. and Dey, N., 2020. *Texture feature extraction techniques for image recognition*. Springer Singapore.
- Chen, G., et al., 2018. Geographic object-based image analysis (GEOBIA): emerging trends and future opportunities. *GIScience & Remote Sensing*, 55 (2), 159–182. doi:[10.1080/15481603.2018.1426092](https://doi.org/10.1080/15481603.2018.1426092)
- Chen, Z., et al., 2019. Corse-to-fine road extraction based on local Dirichlet mixture models and multiscale-high-order deep learning. *IEEE Transactions on Intelligent Transportation Systems*, 21 (10), 4283–4293. doi:[10.1109/TITS.2019.2939536](https://doi.org/10.1109/TITS.2019.2939536)
- Chen, Z., et al., 2022. Road extraction in remote sensing data: a survey. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102833. doi:[10.1016/j.jag.2022.102833](https://doi.org/10.1016/j.jag.2022.102833)
- Dai, L., Zhang, G., and Zhang, R., 2023. RADANet: road augmented deformable attention network for road extraction from complex high-resolution remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 61, 1–13. doi:[10.1109/TGRS.2023.3237561](https://doi.org/10.1109/TGRS.2023.3237561)
- Das, S., Mirnalinee, T.T., and Varghese, K., 2011. Use of salient features for the design of a multistage framework to extract roads from high-resolution multispectral satellite images. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (10), 3906–3931. doi:[10.1109/TGRS.2011.2136381](https://doi.org/10.1109/TGRS.2011.2136381)
- Duda, R.O. and Hart, P.E., 2006. *Pattern classification*. India: John Wiley & Sons.
- Guo, Q. and Wang, Z., 2020. A self-supervised learning framework for road centerline extraction from high-resolution remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 4451–4461. doi:[10.1109/JSTARS.2020.3014242](https://doi.org/10.1109/JSTARS.2020.3014242)
- Gupta, A., Watson, S., and Yin, H., 2021. Deep learning-based aerial image segmentation with open data for disaster impact assessment. *Neurocomputing*, 439, 22–33. doi:[10.1016/j.neucom.2020.02.139](https://doi.org/10.1016/j.neucom.2020.02.139)
- Hu, C., et al., 2018. Model-based segmentation of image data using spatially constrained mixture models. *Neurocomputing*, 283, 214–227. doi:[10.1016/j.neucom.2017.12.033](https://doi.org/10.1016/j.neucom.2017.12.033)
- Kanevski, M., Timonin, V., and Pozdnukhov, A., 2009. *Machine learning for spatial environmental data: theory, applications, and software*. 1st ed. Florida: EPFL Press. doi:[10.1201/9781439808085](https://doi.org/10.1201/9781439808085)
- Kearney, S.P., et al., 2020. Maintaining accurate, current, rural road network data: an extraction and updating routine using RapidEye, participatory GIS, and deep learning. *International Journal of Applied Earth Observation and Geoinformation*, 87, 102031. doi:[10.1016/j.jag.2019.102031](https://doi.org/10.1016/j.jag.2019.102031)
- Kotaridis, I. and Lazaridou, M., 2021. Remote sensing image segmentation advances: a meta-analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 173, 309–322. doi:[10.1016/j.isprsjprs.2021.01.020](https://doi.org/10.1016/j.isprsjprs.2021.01.020)
- Kroese, D.P., et al., 2019. *Data science and machine learning: mathematical and statistical methods*. 1st ed. Florida: Chapman and Hall/CRC. doi:[10.1201/9780367816971](https://doi.org/10.1201/9780367816971)
- Li, X., et al., 2020. Topology-enhanced urban road extraction via a geographic feature-enhanced network. *IEEE Transactions on Geoscience and Remote Sensing*, 58 (12), 8819–8830. doi:[10.1109/TGRS.2020.2991006](https://doi.org/10.1109/TGRS.2020.2991006)
- Lian, R., et al., 2020. Road extraction methods in high-resolution remote sensing images: a comprehensive review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5489–5507. doi:[10.1109/JSTARS.2020.3023549](https://doi.org/10.1109/JSTARS.2020.3023549)
- Li, J., Hu, Q., and Ai, M., 2018. Unsupervised road extraction via a Gaussian mixture model with object-based features. *International Journal of Remote Sensing*, 39 (8), 2421–2440. doi:[10.1080/01431161.2018.1425563](https://doi.org/10.1080/01431161.2018.1425563)
- Liu, R., et al., 2019. A semi-supervised high-level feature selection framework for road centerline extraction. *IEEE Geoscience and Remote Sensing Letters*, 17 (5), 894–898. doi:[10.1109/LGRS.2019.2931928](https://doi.org/10.1109/LGRS.2019.2931928)
- Lu, X., et al., 2019. Multi-scale and multi-task deep learning framework for automatic road extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 57 (11), 9362–9377. doi:[10.1109/TGRS.2019.2926397](https://doi.org/10.1109/TGRS.2019.2926397)
- Maboudi, M., et al., 2017. Object-based road extraction from satellite images using ant colony optimization. *International Journal of Remote Sensing*, 38 (1), 179–198. doi:[10.1080/01431161.2016.1264026](https://doi.org/10.1080/01431161.2016.1264026)
- Maboudi, M., et al., 2018. Integrating fuzzy object based image analysis and ant colony optimization for road extraction from remotely sensed images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 138, 151–163. doi:[10.1016/j.isprsjprs.2017.11.014](https://doi.org/10.1016/j.isprsjprs.2017.11.014)

- Mantripragada, K. and Qureshi, F.Z., 2022. Evaluation of Dirichlet process Gaussian mixtures for segmentation on noisy hyperspectral images. *Computer Vision and Pattern Recognition*, 1–12. arXiv preprint arXiv:2203.02820.
- Martins, V.S., et al., 2020. Exploring multiscale object-based convolutional neural network (multi-OCNN) for remote sensing image classification at high spatial resolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, 168, 56–73. doi:[10.1016/j.isprsjprs.2020.08.004](https://doi.org/10.1016/j.isprsjprs.2020.08.004)
- Mena, J.B. and Malpica, J.A., 2005. An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery. *Pattern Recognition Letters*, 26 (9), 1201–1220. doi:[10.1016/j.patrec.2004.11.005](https://doi.org/10.1016/j.patrec.2004.11.005)
- Miao, Z., et al., 2014. A semi-automatic method for road centerline extraction from VHR images. *IEEE Geoscience and Remote Sensing Letters*, 11 (11), 1856–1860. doi:[10.1109/LGRS.2014.2312000](https://doi.org/10.1109/LGRS.2014.2312000)
- Miao, Z., et al., 2015. An object-based method for road network extraction in VHR satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8 (10), 4853–4862. doi:[10.1109/JSTARS.2015.2443552](https://doi.org/10.1109/JSTARS.2015.2443552)
- Minaee, S., et al., 2021. Image segmentation using deep learning: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. doi:[10.1109/TPAMI.2021.3059968](https://doi.org/10.1109/TPAMI.2021.3059968)
- Mnih, V., 2013. *Machine learning for aerial image labeling*. Canada: University of Toronto.
- Pan, Y., et al., 2018. Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11 (10), 3701–3712. doi:[10.1109/JSTARS.2018.2865528](https://doi.org/10.1109/JSTARS.2018.2865528)
- Ramola, A., Shakya, A.K., and Van Pham, D., 2020. Study of statistical methods for texture analysis and their modern evolutions. *Engineering Reports*, 2 (4), e12149. doi:[10.1002/eng2.12149](https://doi.org/10.1002/eng2.12149)
- Song, W., et al., 2017. Unsupervised PolSAR image classification and segmentation using Dirichlet process mixture model and Markov random fields with similarity measure. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10 (8), 3556–3568. doi:[10.1109/JSTARS.2017.2684301](https://doi.org/10.1109/JSTARS.2017.2684301)
- Soni, P.K., Rajpal, N., and Mehta, R., 2020. Semiautomatic road extraction framework based on shape features and LS-SVM from high-resolution images. *Journal of the Indian Society of Remote Sensing*, 48 (3), 513–524. doi:[10.1007/s12524-019-01077-4](https://doi.org/10.1007/s12524-019-01077-4)
- Valero, S., et al., 2010. Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images. *Pattern Recognition Letters*, 31 (10), 1120–1127. doi:[10.1016/j.patrec.2009.12.018](https://doi.org/10.1016/j.patrec.2009.12.018)
- Wiedemann, C., 2003. External evaluation of road networks. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34 (3/W8), 93–98.
- Yadav, D.P., et al., 2020. Automatic urban road extraction from high resolution satellite data using object based image analysis: a fuzzy classification approach. *Journal of Remote Sensing & GIS*, 9 (1), 279.
- Yang, X., et al., 2019. Road detection and centerline extraction via deep recurrent convolutional neural network U-Net. *IEEE Transactions on Geoscience and Remote Sensing*, 57 (9), 7209–7220. doi:[10.1109/TGRS.2019.2912301](https://doi.org/10.1109/TGRS.2019.2912301)
- Yang, Z., et al., 2022b. TransRoadNet: a novel road extraction method for remote sensing images via combining high-level semantic feature and context. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5.
- Yang, M., Yuan, Y., and Liu, G., 2022a. SDUNet: road extraction via spatial enhanced and densely connected UNet. *Pattern Recognition*, 126, 108549. doi:[10.1016/j.patcog.2022.108549](https://doi.org/10.1016/j.patcog.2022.108549)
- Zhang, Y., Liu, J., and Shen, W., 2022. A review of ensemble learning algorithms used in remote sensing applications. *Applied Sciences*, 12 (17), 8654. doi:[10.3390/app12178654](https://doi.org/10.3390/app12178654)
- Zhang, Z., Liu, Q., and Wang, Y., 2018. Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15 (5), 749–753. doi:[10.1109/LGRS.2018.2802944](https://doi.org/10.1109/LGRS.2018.2802944)
- Zhou, Z.H., 2012. *Ensemble methods: foundations and algorithms*. Florida: CRC press.
- Zhu, C., et al., 2005. The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics. *International Journal of Remote Sensing*, 26 (24), 5493–5508. doi:[10.1080/01431160500300354](https://doi.org/10.1080/01431160500300354)