Computing and Artificial Intelligence 2024, 2(2), 570.
https://doi.org/10.59400/cai.v2i2.570

Article

Clustering data analytics of urban land use for change detection

C. Rajabhushanam

School of Computing, Computer Science Engineering, Bharath Institute of Science and Technology, Bharath Institute of Higher Education and Research, Chennai 600073, India; Rajabhushanamc.cse@bharathuniv.ac.in

CITATION

Rajabhushanam C. Pre-Clustering data analytics of urban land use for change detection. Computing and Artificial Intelligence. 2024; 2(2): 570.
https://doi.org/10.59400/cai.v2i2.570

ARTICLE INFO

Received: 23 April 2024
Accepted: 20 June 2024
Available online: 2 July 2024

Abstract: In this study, the author proposes and details a workflow for the spatial-temporal demarcation of urban areal features in 8 cities of Tamilnadu, India. During the inception phase, functional requirements and non-functional parameters are analyzed and designed, within a suitable pixel area and object-oriented derived paradigm. Land use categories are defined from OpenStreetMap (OSM) related works with the scope of conducting climate change, using multispectral sensors onboard Landsat series. Furthermore, we augment the bands dataset with Spatially Invariant Feature Transform (SIFT), Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-Up Index (NDBI), Leaf Area Index (LAI), and Texture base indices, as a means of spatially integrating auto-covariance to stationarity patterns. In doing so, change detection can be pursuit by scaling up the segmentation of regional/zonal boundaries in a multi-dimensional environment, with the aid of Wide Area Networks (WAN) cluster computers such as the BEOWULF/Google Earth Engine clusters. GeoAnalytical measures are analyzed in the design of local and zonal spatial models (GRID, RASTER, DEM, IMAGE COLLECTION). Finally, multi variate geostatistical works are ensued for precision and recall in predictive data analytics. The author proposes reusing machine learning tools (filtering by attribute-based indexing in PaaS clouds) for pattern recognition and visualization of features and feature collection.

Keywords: distributed computing; HPC; multi-spectral imagery; machine learning; AI; local climate change; zonal analysis; spatial data model

1. Introduction

Using satellite imagery in Google Earth Engine (GEE) such as the omnipresent Landsat 8–9 optical sensors, over a time period of 12 years (2012–2024), with the intent of evaluating geo-based change detection, is the main forte in this research endeavour.

Since urban areas such as Chennai (roughly 5904 sq.km) have depicted refactoring of land use and land cover changes primarily through trend analysis and suitability analysis, we intend to apply geospatial data analytics as an inference engine for effective spatial decision support.

In Tamilnadu, India, urbanisation has been an evolving trend since medieval times to present and this phenomenon can be observed in most town expansion plans, as outlined by the 2011 census. Measuring population growth in Tamilnadu can be attributed to rapid industrialisation in urban areas and the consolidation of cooperative unions from villages to towns.

By generating pre clusters within a region, by unsupervised classification technique, we can assimilate homogeneous feature maps into a class based hierarchical clustering method. This study differs from related literature reviews, in that an instance is defined within a contextual purview, for processing each land-use feature by deducing in a neighbourhood distance-based classifier.

2. Objective

Demarcating zonal boundary or distinct urban gradient break is the primary forte of optical remote sensing sensors during the present day. In this study we intend to demonstrate the findings from one such controlled experiment, which is to provide to government authorities a spatial decision support system (SDSS).

Primary focus is in geospatial bigdata analytics and its application to cadastral mapping in urban zones of Tamilnadu. Specifically, we will address the following:

- Implementing a base layer for Tamilnadu state with pre-determined feature classes.
- Visualization of topology conditions for any given pixel in the base geo-referenced data.
- Overlay of raster feature datasets with vector feature classes for regular/irregular spatial delineations.
- Performing spatial queries with selection, projection, and joins to extract a geospatial feature map for the geoanalytical engine using Google Earth Engine (GEE).
- Implementing a spatial-temporal (from 2012 to 2024) inference model using machine learning and Platform as A Service (PaaS) cloud offering. The linear model simulates climatic and aspatial change detection techniques, within a linear multivariate geostatistical regression criterion.
- The model is trained with region of interest (ROI) and point of interest (POI) sample datasets, and then it is applied for validating feature datasets from any given feature map, using iterator operator and collection classes in JavaScript programming language.
- Data analytics is applied to operate on the following geospatial change detection in LANDSAT-8, LANDSAT-9 sensor's-based satellite imagery scenes.
- Trend analysis with focal and zonal areal interpolation boundaries.
- Suitability analysis with environmental and human induced factors.
- Unsupervised classification study is then under-taken to do density clustering of data points, resulting in user specific intervals for each user-specific class. Deriving overall accuracy, specificity, and recall measures from a classification study. Specifically, a confusion matrix (commission and omission errors) is computed.
- Furthermore, Spatially Invariant Feature Transform (SIFT) method is applied to the feature map and the classification study is ensued for the selected multi-scale scene analysis. True positives and False negatives are computed and the root mean square error (RMSE) is conducted for validation and verification.
- Finally, Quantile regression analysis is plotted using Mean-Covariance shift measure as a planar separator to fit to the datapoints. Slope and Intercept are calculated to determine the inflection point (feature) where collinearity is evident.

3. Related study

Shan and Sampath [1] proposed a method based on a region-growing algorithm to group similar points into the same building by iteratively collecting points within a moving window. In the study of Sampath and Shan [2], the fuzzy c-means (fuzzy k-means) algorithm was used to cluster individual planar roof segments to reconstruct building models. As the k-means and fuzzy c-means algorithms can only cluster convex shapes, they are rarely utilized directly to segment individual building derived from image segmentation [3–6].

The task of retrieving parameters from observations/predictions is most commonly associated with deriving measured output. With this research proposal, we aim in creating an inverse radiometric calibration machine learning model. This model will map causal inference from live sensor readings to observed bio-geophysical readings. This approach varies from training models in Artificial Intelligence and Machine learning domains [7]. Further, the feature extraction task can be based on multi-dimensional sensor measurements or the target parameter can consist of several variables. Thus, this study is categorized as a classic Classification and Regression Theory (CART) problem.

Our hypothesis is that linear classification technique produces spatially invariant features within a zonal areal modifiable area unit problem (MAUP). In doing so, adapting to multiple perceptron neural networks, transforms inputs to expected outputs by iteratively clustering over bounded feature maps (DBSCAN-VAE). So, as a result non-linear methods such as NN, can be deployed in a generative network (hyper-parameters) in a linear mode. The use of neural autoencoder allowed us to transform raw images in their feature representation versions. These encoded images contain all the essential information about the original. Among different neural network models, autoencoders have found the application in many domains. In image processing, the autoencoders are used mostly for feature extraction, image segmentation, image compression/data reprojection and image reconstruction. The advantage of our autoencoder is that it can be used as a standalone algorithm for feature extraction with further use of extracted features for unsupervised purposes.

As this study is a confluence of computer-vision, remote sensing, machine learning, and deep learning fields, a select number of publications have explored the research criterion. By spatially interpolating with deep learning approaches, such as spatial-temporal VAEs, allow learning the behaviour of simulations from previous runs. Ideally, this results in speeding up climate models by regularization of latent factors. While there will be a penalty on prediction accuracy, the speed-up might be worth in some cases. In addition, the same unsupervised autoencoder-like models could be used to approximate the sub grid processes Such architectures allow for the stochastic modelling of the earth system from raw observations. Currently, no use of this approach for the data-driven modelling of earth dynamics has been published.

4. Synthesis of feature maps by using a variational autoencoder (VAE)

The idea behind a VAE is different from the idea of an Autoencoder. Instead of reproducing single inputs, the goal is to map the unknown distribution of the inputs onto a d-dimensional multivariate Gaussian distribution and then denormalize to original data (**Figure 1**). The representation of the latent space using a dimensionality reduction method is also explored in the study of Lu et al. [8]. The goal was to detect and track extreme events, using an AE as a method to extract features.

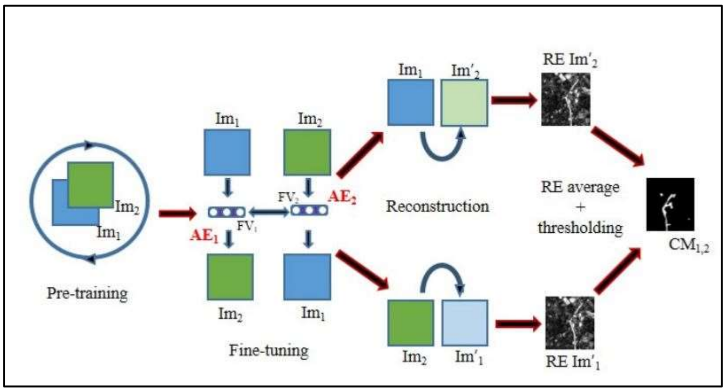

Figure 1. Schematics of an auto-encoder in image synthesis [9].

Adapting the effectiveness of autoencoders to such a study is particularly important, as the final output is a generative model that closely approximates the latent space with second order derivative (*dy/dx*). Autoencoders can better represent complex spatial-temporal nature of the underlying datasets that is currently often summarized with indices or convoluted features. In particular there are usually no labels making unsupervised autoencoders useful. The motivation for AEs comes from: 1) the unsupervised extraction of useful features to predict; 2) their ability to generate additive (unseen data) samples from the original distribution; 3) the removal of noise, and 4) the identification of anomalies.

In the worst case, such as without the use of an accelerating index structure (applying a sequential search), or on a degenerated dataset (e.g., all points within a distance less than e), the overall runtime complexity remains $O(n^2)$. Furthermore, the time complexity for creating spatial index structures (e.g., GiST or R*-tree) is usually in $O(n\log n)$ and, as a result, this does not affect the total time complexity. Although some border points may be firstly marked as outliers and be clustered later, they are not added to the seeds list and unnecessary neighbourhood retrieval avoided because they are not core points. On the other hand, a border point may be shared by two closely distributed clusters. In this case, the algorithm will group the shared point into the first discovered cluster. An improvement on this situation is to group it into the cluster that its nearest core point belongs. Except from these cases, the result is insensitive to the processing order of the points. Additionally, if treating border points as outliers, the algorithm is deterministic.

Previous studies on land use mostly focused on the functional zoning of ecological landscapes or farmland, with limitations to other implementations [10]. As mentioned above, OSM data contain multiple and finer information on land use, especially social functional information that is rarely captured from remote sensing images. This study pays attention to urban social functional land use and tries to extract three social functional types of land use based on OSM data. The OSM database contains 21 features describing the land surface objects, and each feature is labelled by several attributes. According to the mapping rules of **Table 1**, this study chooses three land use function types as follows: 1) residential land use, which indicates the land parcels that are labelled ‘residential’, ‘apartments’, ‘dormitory’, ‘garage’, ‘house’, ‘residential’, ‘dorm’, etc.; 2) commercial land use, which refers to the land for wholesale and retail, accommodation and catering, and commercial or financial purposes with labels of ‘retail’. ‘Marketplace,’ ‘pharmacy’, ‘hotel’, ‘café’, ‘restaurant’, ‘bank’, ‘commercial’, etc.; 3) public service land use, which includes government and organizations, science and education, medical and health charity, recreational land, public facilities, park and green space, and scenic facilities with labels of ‘office’, ‘university’, ‘hospital’, ‘cinema’,’ post office’, ‘park’, ‘viewpoint’.

Table 1. Confluence of image analysis techniques.

S.NO	Scientific domain	Technology	Roadmap
1	Distributed Computing	Parallel and Distributed High Performance Computing System (HPCS).	High level design (HLD) to integrate satellite imagery retrieval, storage, and analysis into a value-added reseller (VAR) functional outlay.
2	Computer Vision	Algorithm Development for Scheduling and Staging System	Image analysis, Digital Image processing for pixel detection and synthesis. Includes 3-D (Stereo vision) and 2-D feature extraction.
3	Satellite Communications	Image Retrieval and Archival System	Building Block for Ingest and Analysis workflows.
4	Geospatial Analytics	Feature Engineering	Geoprocessing, Recomputing, Geoanalytical, and Geo-Statistics.
5	Cloud Computing	Client/Server n-Tier Architecture in PaaS and SaaS.	NIST defined a computer based functional entity with Ubiquity and Pervasiveness.
6	Remote Sensing	Imaging System for Sensing and Remote Measurement of natural phenomena	Radiometric measurement and feature extraction
7	Software Engineering	JavaScript and Python software development using Google Earth Engine (GEE) as an Interface (ICD standards)	Modular driven architecture with call-return graphs and tight coupling using high cohesiveness.
8	Machine Learning	Unsupervised Classifier, Variational Autoencoder, SGD and Cross-Entropy feature segmentation.	Semi-structured data analysis using unsupervised learning within a pre-defined number of feature classes.
9	Artificial Intelligence	Inference System with Explainable Artificial Intelligence (XAI)	Deep-learning using feedback and reward inference engine.
10	Database Technologies	SQL data store for benchmarking and Indexing purpose.	Data store in the cloud and within the firewall premises for maintaining redundancy and replication ability.

5. Proposed methodology

Deep learning also faces challenges that are unique to earth science data: multimodality; high degree of heterogeneity in space and time; and the fact that earth science data can only provide an incomplete and noisy view of the underlying eco-geo-physical processes that are interacting and unfolding at different spatial and temporal scales.

Causal reasoning with an inference engine is limited to studies for one site or one instance by design. To generalize the results obtained as in a sequential work flow, we have to aggregate the effects from each normalized input as a feedback mechanism to the precedent and antecedent path. Then applying heuristic rules (knowledge base aided by a human in the loop) we can reconstruct any input sequence averaging over mean (zero) and variance (absolute one) [10,11].

In this paper, a density-based cluster algorithm, DBSCAN [12], is independently implemented in a spatial database to separate building point clouds into individual buildings. As a density-based clustering method, DBSCAN characterizes a well-defined "density reachability" cluster model by connecting points that satisfy a density criterion as defined as a minimum number of objects within a certain neighbourhood.

In contrast to k-means method that can only find convex clusters, DBSCAN can form a cluster of an arbitrary shape. Moreover, based on spatial databases, it benefits from the spatial index offered by the system and achieves performance improvements. The algorithm is implemented as follows, and the flowchart is shown in **Figure 2**.

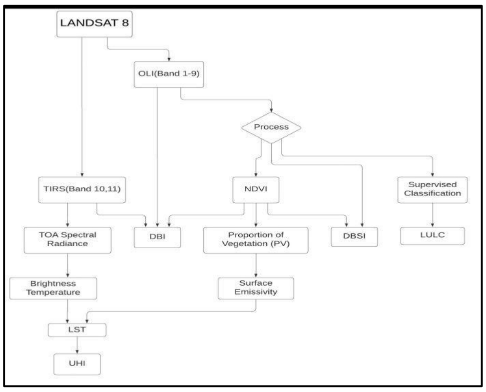

Figure 2. Proposed methodology. NDVI (Normalized difference vegetation index; Dry Built-up index (DBI); Dry bare-soil index (DBSI).

6. DBSCAN algorithm

We propose a variant of the existing DBSCAN algorithm for these purposes. The algorithm efficiency is measured by the relative weights applied to form clusters from a randomly assigned seed value. In addition to the distance (radii) parameter the pre-processing parameter, number of clusters (region of interest—ROI), is defined in a linear approach. For each assigned point in a cluster, intra-class variance is minimized or, we can emphasize that connected component segments are created within a 5 × 5 moving pixel neighbourhood. Thus, each individual data point in dense regions are core points that are assigned to each cluster based on the specificity of the sparse density matrix.

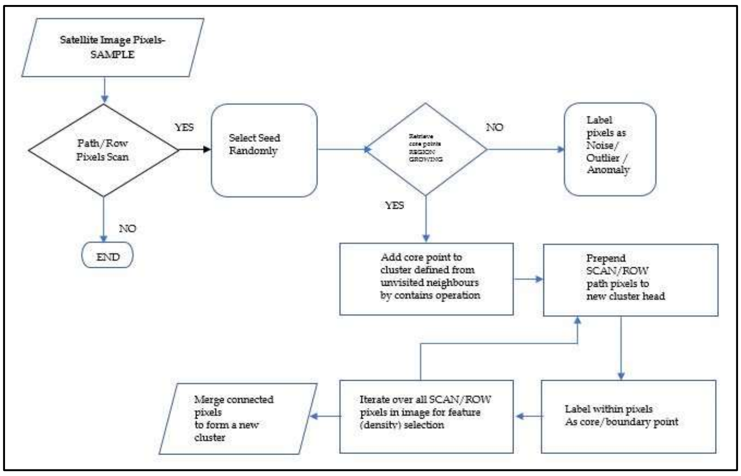

Figure 3. Process data flow diagram in pixel-based approach.

One example of a bio-geophysical phenomena is surface/atmospheric temperature (Band 10—TIR band of LANDSAT-8) recorded from the satellite, is input to the climate forecasting model. This topic is sparsely researched in the remote sensing literature, and hence we emphasize the value factor derived by conducting a continuous multi-variate as inputs to our study. Many other input variables can be listed—NDVI, DBI, DBSI, etc. (**Figure 3**).

In this experiment pre-processing machine learning model VGG_{16} is adapted (transfer learning) for filtering patterns during the training. During this iteration step input layer, hidden layers, and output layer are connected in a sequential arrangement to form a non-singularity transfer learning entity (**Figure 4**).

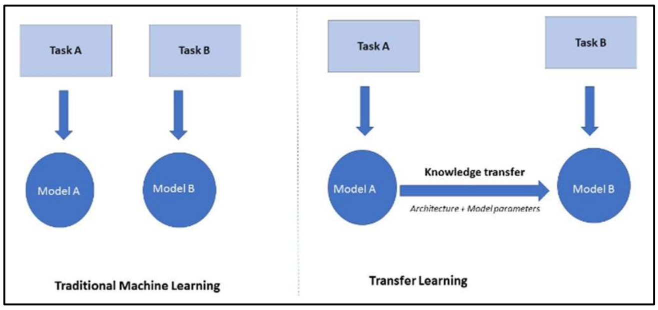

Figure 4. Type I and II error reduction.

The filters perform feature extraction and transform our data into a latent space where the problem is potentially easier to solve. For VGG_{16}, data is typically structured in 2-D space (Longitude, Latitude). We can do data augmentation by inputting the sequence of bands (channels) and Indices from satellite data into the feature selection model (**Figure 5**). This architecture is akin to a convolutional neural net, excepting that it behaves in a VGG_{16} pre-trained network as such. For our purposes, we shall denote this step as Model-A. Inputs to the variational autoencoder (VAE) are denoted as Model-B.

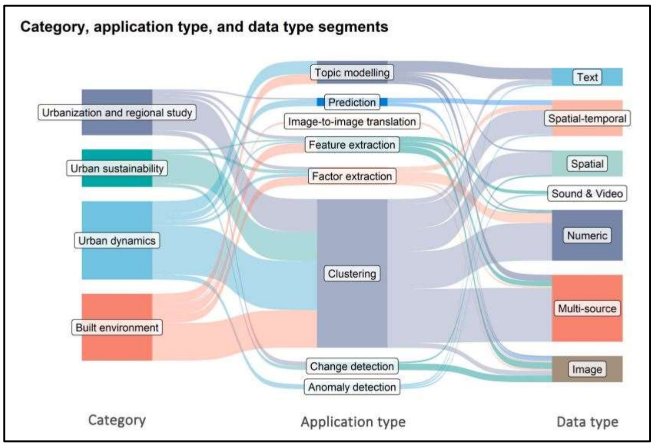

Figure 5. Category, application type, and data type segments [13].

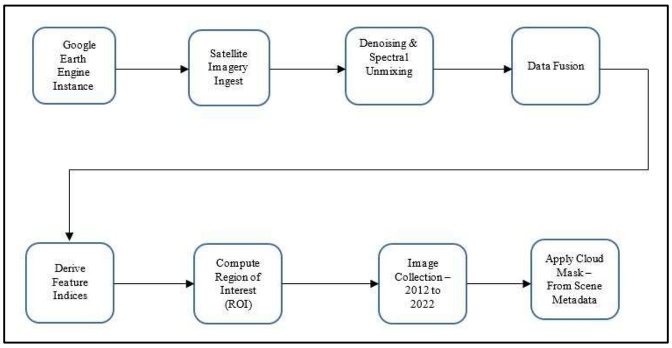

Figure 6. Pre-processing of input imagery. Examines an instance of ROI.

Therefore, we can adapt a unified representation with x = height, y = width, b = band, k-statistical indices, and T-Temporal identity matrix (**Figure 6**). This would result in cubic filters with T channels. Two ways of modelling spatial, spectral and temporal relationships. (a) Cubic convolutions over space and frequencies and stacking the n time steps. (b) Stacking frequency, B, and time, T, and performing 2D convolutions over spatial dimensions (X, Y). Originally, the concept was developed for visual recognition problems with time-varying inputs [14]. It is important to note that when stacking channels, e.g. (B, T), the order of them is not taken into account by the model (**Figure 7**).

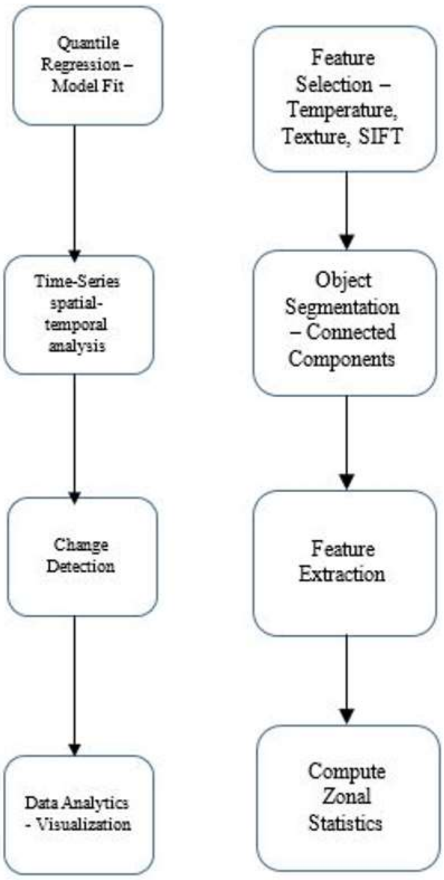

Figure 7. Post-processing flow chart.

Mean Square Error or Cross-entropy Optimization?

Retrieval (**Figure 8**) is most often associated with least-square regression modelling. Traditionally, linear regressors optimized by least-square have been the preferred choice but are not always sufficient to capture the complexity of retrieval problems. A neural network optimized by the mean-square error (MSE) loss function, is a one way of extending the least square linear regression with non-linearity. Alternatively, a probability distribution over possible outcomes can be modelled with Cross-Entropy based error functions. Cross-Entropy is generally associated with problems where we wish to label data, e.g., segmentation or classification. It can be shown [15] that the MSE loss is the maximum likelihood solution to a problem where the target can take any real values.

Many challenges exist for deep learning in bio-geophysical parameter retrieval problems. Since we are modelling the Earth's state, we need to apply algorithms on large amounts of data. Further, we have many sources of variance in our observations caused by e.g., seasonal, yearly and geographical variation (**Table 1**).

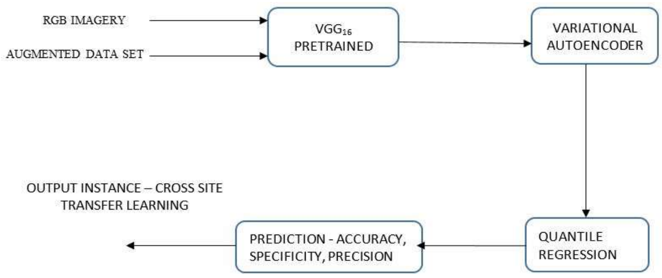

Figure 8. Illustration of a generalized network for scene understanding and refinement.

6.1. Dataset preparation

Landsat 8, 9 sensor-based satellite imagery will be ingested from Google earth engine (GEE) together in creating a spatial data collection service. Bands in near-infrared, short-wave infrared, thermal, and panchromatic will be chosen as inputs to the pre-trained network VGG_{16}.

From the chosen band spectrum, indices will be derived for first—Creating an image fusion-based collection set. Namely panchromatic sensor from Landsat-8 at 15 m resolution will be fused with 30 m nir-sw-tir bands to create a spatial dataset that will reveal distinct patterns for a region of study (**Table 2**).

Table 2. Specification of image analysis domains.

S. No	Geospatial domain	Description
1	Geo-Processing	Zonal statistics, Feature dataset, Feature understanding
2	Geo-Computing	Feature selection, Feature extraction, Attribute spatial joins, Model fitting and Feature understanding
3	Geo-Analytical	Model fitting, Validation, Predictive analytics
4	Geo-Statistics	Validation, Predictive analytics

After the image fusion step, spatial indices (**Figure 9**) calculated from each pixel in the input satellite imagery will be computed as inputs to the pre-trained VGG_{16} network. Specifically:

- Normalized Difference Vegetation Index (NDVI) for Vegetation feature class.
- Built -up Dryland—Dry Built-Up Index (DBI).

- Barren land in dry climates—Dry Bare Soil Index (DBSI).
- Normalized Difference Vegetation Index (NDVI) for water bodies feature class.

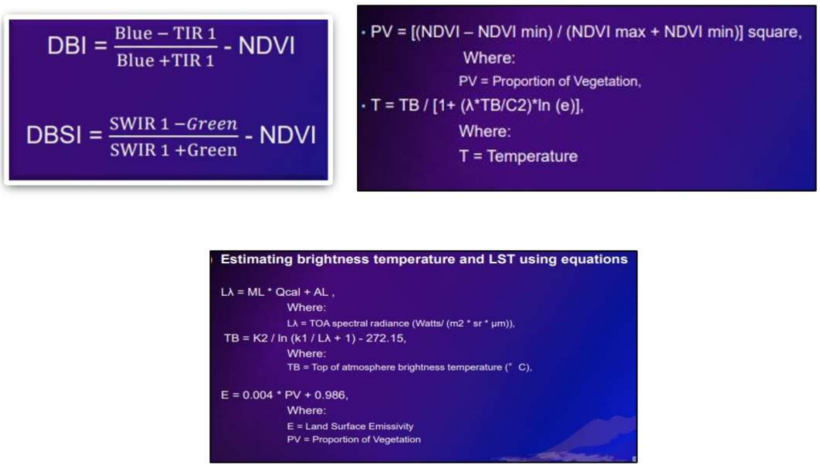

Figure 9. Spectral indices calculation using image expression.

6.2. Data augmentation in input feature dataset

- Scale Invariant Feature Transform (SIFT)—Hierarchical Scale Space (HSS).
- Texture derived from Gray Level Co-Occurrences Matrix (GLCM)-entropy.
- Temperature feature derived from Band 10 of Landsat-8,9 (**Table 3**).

Table 3. Summary of geo-processing tasks.

S.No	TASK—Unit Instance	Description
1	Zonal Statistics	Deriving exploratory statistics for Region of Interest (ROI)
2	Feature Selection	For pattern recognition and discriminative analysis
3	Feature Dataset	Multi-Variate inputs and Derived classes in multi-dimensional analysis.
4	Feature Extraction	Fuzzy Inference System based Segmentation followed by CART task.
5	Feature Understanding	Variational Autoencoder (VAE) as a generative adversarial network for encoding and decoding patterns that can be discernible by a human in the loop.
6	Attribute based Spatial Joins	Linking an object-relational data store with Geospatial data using spatial joins for query and search operations.
7	Model Fitting	Quantile Regression technique for Interval demarcation and polynomial expansion using a symmetrical estimation.
8	Validation and Verification (V&V)	Conducting n-fold cross validation to achieve conditional local minima for ROC characteristics.
9	Predictive Analytics	Deep Learning technique to generate specificity, accuracy, and precision for recall instance.

7. Algorithm

Algorithm 1 Regularization and hyperparameters tuning

1: Aim is to achieve local minima through density clustering (DBSCAN) based on Euclidean distance measure.
2: Design covariance matrix as X^T.
3: Design of seed value as a random variable (RV).
4: For each seed, initialize core point, boundary point and zonal area (ROI).
5: While region growing is not ~ to NIL for batch intervals while computing weight/density estimate.
6: Initialize cluster radius to be a compact shape (necessarily not rectangle, square) CIRCULAR path.

Algorithm 1 (*Continued*)	
7:	Initialize region growing pixel values that in time and space will merge all mutually exclusive points.
8:	If core point distance is LT maximum points, assign boundary.
9:	If cluster point distance GT minimum points, assign cluster region.
10:	Now, repeat over all pixels in the sample (~) for consideration into neighbourhood (optimal) having low intraclass variance.
11:	Construct covariance matrix.
12:	Stop when number of cluster regions reach 4 feature classes.
13:	For any unassigned point, assign it as an outlier/anomaly.
14:	Stop region growing.
15:	Scan all boundary points € cluster regions. If shape is defined as equidistant, assign count of core points as a threshold.
16:	Finally run R-Tree indexing for efficient spatial index retrieval.

8. Design of input parameters to the model

- Design of a sparse matrix;
- Design of covariance matrix;
- Design of features as inputs to VAE—pyramidal resolution;
- Design of spatial autocovariance measure;
- Design of input tiling scheme in imagery;
- Design of UTM zone in region of study (ROI);
- Design of thresholding value in normalization of input data;
- Design of SIGMA value as input to SIFT;
- Design of cell size during Texture calculation with GLCM method;
- Design of hidden layers in variational autoencoder (VAE);
- Design of sequence flow during input to pre-trained network (VGG_{16});
- Design of ordering of inputs to VAE;
- Design of radius parameter for input to density clustering;
- Design of neighbourhood size (5 × 5 default);
- Design of epoch size during sequential processing in Iterator to avoid over/under fitting.

9. Conclusion

In this study, feature engineering paradigm was applied to evince interest and focus on the fragile relationships between humans and factors in environmental eco-system. We have dealt with multi-spectral sensors and the spatial scale problem, in achieving an optimal scale-resolution under which causal inference is evident. Various state of the art algorithms and pixel/object-oriented machine learning stages were feature selected and feature extracted, in accordance to spatially distributed parallel systems.

Conflict of interest: The author declares no conflict of interest.

References

1. Shan J, Sampath A. Urban Terrain and Building Extraction from Airborne LIDAR Data. CRC Press; 2006.
2. Sampath A, Shan J. Segmentation and Reconstruction of Polyhedral Building Roofs From Aerial Lidar Point Clouds. IEEE

Transactions on Geoscience and Remote Sensing. 2010; 48(3): 1554-1567. doi: 10.1109/tgrs.2009.2030180
3. Awrangjeb M, Fraser C. Automatic Segmentation of Raw LIDAR Data for Extraction of Building Roofs. Remote Sensing. 2014; 6(5): 3716-3751. doi: 10.3390/rs6053716
4. Caihua Y, Yonghong L, Weijun Q, et al. Application of Urban Thermal Environment Monitoring Based on Remote Sensing in Beijing. Procedia Environmental Sciences. 2011; 11: 1424-1433. doi: 10.1016/j.proenv.2011.12.214
5. Sohn G, Dowman I. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing. 2007; 62(1): 43-63. doi: 10.1016/j.isprsjprs.2007.01.001
6. Zhou H, Jiang H, Zhou G, et al. Monitoring the change of urban wetland using high spatial resolution remote sensing data. International Journal of Remote Sensing. 2010; 31(7): 1717-1731. doi: 10.1080/01431160902926608
7. Tsagkatakis G, Aidini A, Fotiadou K, et al. Survey of Deep-Learning Approaches for Remote Sensing Observation Enhancement. Sensors. 2019; 19(18): 3929. doi: 10.3390/s19183929
8. Lu X, Yuan Y, Zheng X. Joint Dictionary Learning for Multispectral Change Detection. IEEE Transactions on Cybernetics. 2017; 47(4): 884-897. doi: 10.1109/tcyb.2016.2531179
9. Sublime J, Kalinicheva E. Automatic Post-Disaster Damage Mapping Using Deep-Learning Techniques for Change Detection: Case Study of the Tohoku Tsunami. Remote Sensing. 2019; 11(9): 1123. doi: 10.3390/rs11091123
10. Schäfer M, Kröger M. Joint problem framing in sustainable land use research. Land Use Policy. 2016; 57: 526-539. doi: 10.1016/j.landusepol.2016.06.013
11. Tibau XA, Reimers C, Eyring V, et al. Spatiotemporal model for benchmarking causal discovery algorithms. Published online March 23, 2020. Doi: 10.5194/egusphere-egu2020-9604
12. Bengio Y, Deleu T, Rahaman N, et al. A Meta-Transfer Objective for Learning to Disentangle Causal Mechanisms. ArXiv. 2019; arXiv:1901.10912v2.
13. Wang J, Biljecki F. Unsupervised machine learning in urban studies: A systematic review of applications. Cities. 2022; 129: 103925. doi: 10.1016/j.cities.2022.103925
14. Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. KDD. 1996; 96(34): 226-231.
15. Donahue J, Hendricks LA, Guadarrama S, et al. Long-term recurrent convolutional networks for visual recognition and description. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).