LAND USE AND LAND COVER MAPPING USING TIME SERIES DATA

V.Yaswanth[1], V.Hari Sudarsanam[2], Suneetha Manne[3]

[1,2,3] Information Technology, VRSEC, Vijayawada, Andhra Pradesh, India.
varadayaswanth@gmail.com
harisudarsanam2@gmail.com
suneethamanne74@gmail.com

Abstract:

The detection of changes in land use and land cover is a crucial task in the monitoring and management of natural resources. In this instance, we differentiate two common approaches to detect changes over time: Convolutional Neural Network (CNN) and Time-Weighted Dynamic Time Warping (TWDTW) algorithms. We assessed the attainment of both methods on two real data sets of MODIS satellite imagery and compared their accuracy over days. According to findings, while both methods can accurately detect changes, TWDTW outperforms CNN in detecting gradual changes that take place over long periods of time. We also examined the computational efficiency of both methods and found that CNN is faster than TWDTW, making it more suitable for real-time applications. Our results provide insight into the strengths and limitations of these two methods for the ability to comprehend trends. Combining these two methods can provide a more comprehensive approach to accurately detect alterations in land usage over time. The results of this instance may influence the selection of appropriate methods for specific applications and provide a basis for future research in this area.

Keywords: MODIS Satellite images, Land Use Land Cover (LULC), Convolutional Neural Network (CNN) Algorithm and TWDTW algorithm.

1. Introduction:

Most of major environment alterations that occur at the scale of landscapes and directly affect ecosystem processes include industrialization, deforestation, and agricultural development. Biological interactions, however, can also alter the physical makeup of landscapes and create spatial patterns within them. Ecosystem engineering is the name given to this phenomenon. It is essential to map the LCLU at the landscape size in order to monitor and manage these changes. Initiating this effort with classification using satellite data is a crucial step. The material that coats the surface of the earth, such as vegetation, bare soil, water, and urban infrastructure, is referred to as “land cover”. The estimation of land cover generates the fundamental information needed to carry out tasks like thematic mapping and machine learning based investigations. The descriptor “land use” reflects how a portion of land is exploited, such as for amusement, agriculture, or as a sanctuary for wildlife. The land cover is the actual substance that makes up the earth's surface. How humans use the land for socioeconomic activities, this is renowned as "land usage". The layout of human operations and environmental phenomena on the ecosystems over a certain time using prominent biological and statistical methods of analysis of pertinent source data is typically a prerequisite when the filler words "Land Use" and "Land Cover" are combined.

Contribution:

```
df_weight <- data.frame(Difference, Logistic = log_fun(Difference,10), Linear = lin_fun(Difference,10))
```

The above line of code shows the building of a data-frame by combining the Linear and Logistic time-weight models of TW-DTW algorithm.
A hybrid model was built by combining linear model and logistic model in TW-DTW algorithm that helps in improving the efficiency of the result that is obtained when compared the results obtained by linear model and logistic model individually.

2. Object-based Classification:

Prior to individually distinguishing each object, object-based classification technique associate image pixels using a features extraction procedure to generate spatial frequency cohesive visual features, subjected to pixel-based tagging methods that designate each pixel peculiarly. The step entails defining pixels according to their spectral data, form, glossiness, and regional distribution with neighbouring pixels. Contrary to even more proven, pixel-based approaches, these solutions were introduced recently. Pixel-based system is greatly dependent on spectroscopy, present in each individual pixel, as opposed to object-based classification; it makes use of knowledge from a set of related pixels known as object classes collectively. Pattern objects or features are aggregates of pixels that overlap spectral characteristics like dimensions, shapes, and background from the pixels vicinity. This is a 2-step process which begins with segmenting the image into discrete items or characteristics, then classifying each object separately. This classification is an effort to mimic the type of analysis that viewers employ to comprehend pictures.

3. Multi-Spectral Images:

Multispectral imaging is the name given to spectral imaging methods that provide images that correspond to at least a few spectrum channels, and sometimes as many as ten. The commonly used spectral regions frequently extend beyond of the spectrum of visible light, or span elements like UV as well as infrared spectrum. Although though multispectral imagers frequently provide wavelength channels that don't overlap with a greater number than three channels, color photography might theoretically be used as a multispectral imaging technique. Economic environments contain both imaging equipment and finger gadgets. Confocal sensors, for instance, are regularly altered for specific applications, especially when it comes to of the spectral bands involved. Space networks and avionics use high resolution laser scanners, commonly referred to as modis lenses.

4. Land Use and Land Cover:

The characterization of human endeavors and environmental features on the surface throughout a specific period of time using accepted scientific as well as statistical techniques for the examination of relevant original sources is known as LULC. It can be categorized in numerous ways. LULC components can take on many different shapes, including built-up or Land types include forest, farming, and urban areas. LULC maps can be utilized for a numerous purposes, including the administration of natural resources, benchmark tracing for GIS input, establishing legal bounds for tax and estate investigation, and many more. It is impossible to map LULC without the use of additional geographic datasets. Customers may find out about the financial paradigm using maps of a region's land use as well as its land cover (LULC). The continuous tracking of the dynamic nature of agro ecosystems, forested restorations, bodies of water at the surface, and other phenomena will be possible thanks to the incorporation of LULC statistics in nationwide large datasets.

5. Architecture Diagram:

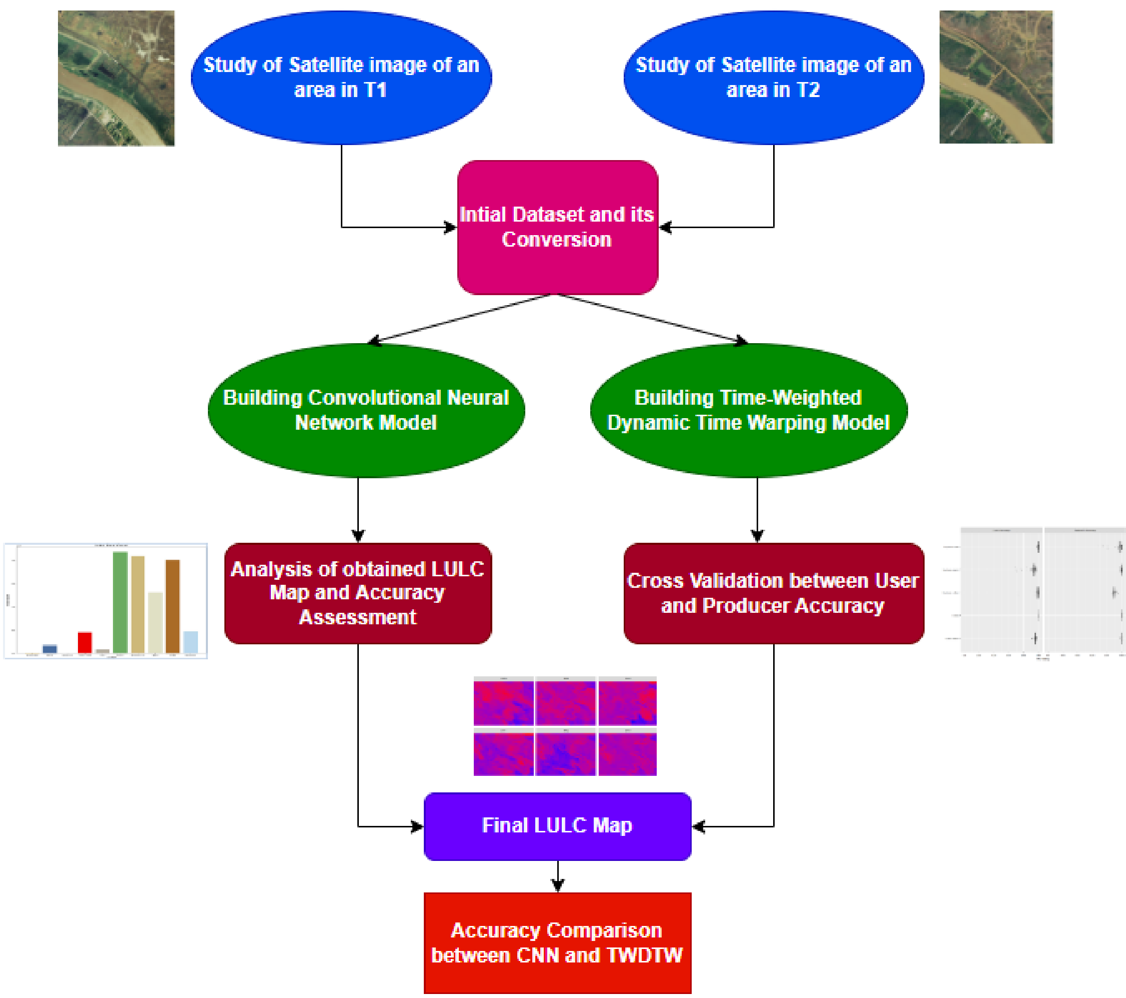

Fig-1: System Architecture Diagram

Fig-1 that is presented below depicts the project's working model. The initial action is to compile a dataset of MODIS satellite images with a range of temporal profiles, of vegetation in South America's Brazilian Amazon. We start by transforming the obtained dataset to Data frames, a way for summarizing visual feature data objectively. Its central idea is to partition the ambient occlusion zone into areas that might be related to base-layer training that is pertinent to the provided instance. The development of two training models, is finished. The initial LULC maps from the two built models are then evaluated.

6. Methodology:

6.1 Dataset Description:

The MOD13Q1 dataset contains 16-day Determinants of habitat from the (MODIS) satellite signal. It offers information for the years 2000 to the present on NDVI and EVI with a locational eloquence of 250 meters. The red and NIR bands of MODIS Data, which are sensitive to vegetation greenness and leaf area index, respectively, are used to calculate the NDVI and EVI. The EVI is intended to be more sensitive to vegetation changes in places with high amounts of atmospheric aerosols or surface reflectance, although the NDVI is frequently employed to monitor vegetation growth and health. The MOD13Q1 dataset is frequently used in many different applications, including ecosystem modeling, crop monitoring, and land cover mapping. It is accessible on the NASA Earth data website and can be accessed and downloaded using a number of programs, including the MODIS Data Download and Visualization Portal (MODIS DVP).

6.2 Algorithms Usage:

6.2.1 Convolutional Neural Network:

A type of deep neural network learning known as CNN is widely employed in industries for charting crop production. CNN consist of 5 layers namely; input layer, convolutional layer, pooling layer, fully-connected layer and output layer.

Table-1: Convolutional Neural Network (CNN) Algorithm

- Assemble and prepare aerial or satellite images.
- Create the fully connected, pooling, along with convolutional layers for the CNN framework and adding filters to the data being provided will enable convolutions.
- Use a ReLU or other activation function to add nonlinearity.

  $$f(s_i) = \frac{1}{1 + e^{-s_i}}$$ ………………………………… **(Sigmoid Activation Function)**

  $$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}}$$ ………………………………… **(Softmax Activation Function)**

  where sj are the scores that the net infers for each and every class in C.
- To decrease the dimensions of space, use pooling techniques (like max pooling).
- Flatten feature maps and join them to layers that are joined completely.
- Give the neurons in the completely linked layers activation functions (like ReLU, for instance).

- Incorporate the quantity of land use and land cover groups into the output layer design.
- Create a suitable loss function, like categorical cross-entropy.

  $CE = -\sum_{i}^{C} t_i log(s_i)$ ……………………………………………… **(Cross-Entropy loss)**

  where the ground-truth for every single class i in C and the CNN score are ti and si, respectively.
- Apply methods for optimization like SGD or Adam to the model's training.
- A set of validation data or cross-validation should be used to assess the model.
- Using newly collected data, apply the model that was trained to anticipate land use and land cover classes.

6.2.2 Time-Weighted Dynamic Time Warping:

Maps can be created using the twdtw algorithm to scrutinize time-series data and classify. The variant of conventional DTW technique that takes the temporal dynamics of data patterns into consideration.

Table-2: TW-DTW Algorithm

- Assemble historical data in series showing the changes in land usage or land cover over the course of time.
- To guarantee uniform scaling, normalize or standardize the time series data.
- Create a function called weighting that gives the variations in associated time series values weights. Gaussian, exponential, and linear functions are frequently used as weighting factors.
- Create a cost matrix to depict the differences between pairs of time series data items. It is computed that the cost matrix C(i, j) = |x(i) - y(j)| * w(|i - j|)
- Use dynamic programming to determine the best alignment path between the two time series and create a cumulative cost matrix D, where D(i, j) is the smallest cumulative cost to go from point (0, 0) to (i, j).

  D(i, j) = C(i, j) + min(D(i-1, j), D(i-1, j-1), D(i-1, j-1)) yields the cumulative cost matrix D.
- To acquire the alignment of the two time series, follow the most advantageous path back via the cumulative cost matrix. To get to the top-left cell, start at the bottom-right cell and go along the path with the least amount of accrued cost.
- Determine the TWDTW distance, which illustrates how the two time series differ from one another.
- In order to establish if the two time series are comparable or dissimilar, apply a threshold to the TWDTW distance.
- Getting the change map, which shows the difference for the selected region of interest between two time lapses, is the last step.

$$\omega_{i,j} = g(t_i, t_j) \qquad (1)$$

$$\omega_{i,j} = \frac{1}{1 + e^{-\alpha(g(t_i,t_j)-\beta)}} \quad (2)$$

Equations (1) and (2) describes the calculation of Linear DTW model and Logistic DTW model where the number of days that occurred across the dates ti in the sequence with tj among the time frame series is denoted by the symbol g(ti, tj).

7. Performance Evaluation and Results:

Confusion Matrices are obtained for chosen area of interest that is used to calculate the performance, visualizing the values obtained and summarizing by implementation of Convolutional Neural Network (CNN) algorithm.

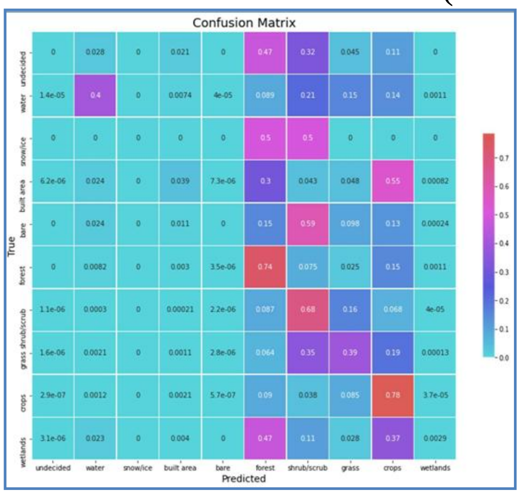

Fig-2a: Confusion Matrix (2016)

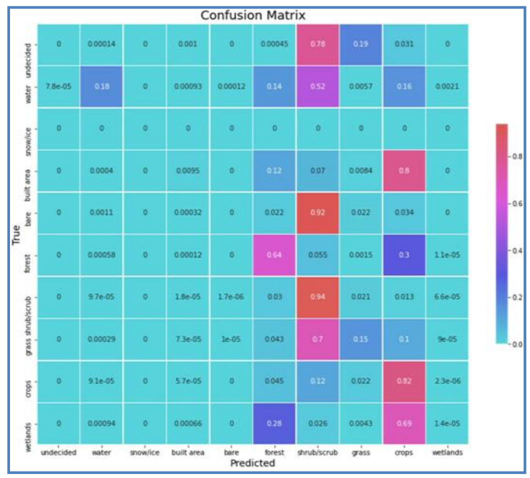

Fig-2b: Confusion Matrix (2021)

Fig-2a, Fig-2b represents the confusion matrices that occurred after the execution of Convolutional Neural Network (CNN) algorithm for the years 2016 and 2021. Change in values was observed when comparing the confusion matrices of the years 2016 and 2021.

Fig-3a: Original Image in 2021

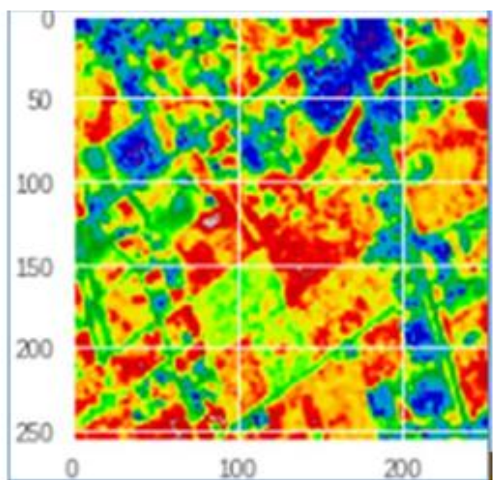

Fig-3b: Resultant Image in 2021

Fig-3a shows the original MODIS Satellite image that was captured in the year 2021 whereas Fig-3b shows the original MODIS Satellite image in 2021 after the classification i.e., resultant image by assigning different colors to different regions like dark blue for water bodies, green for grasslands, red for forest regions and orange color with unclassified region.

Fig-4a: Original Image in 2016

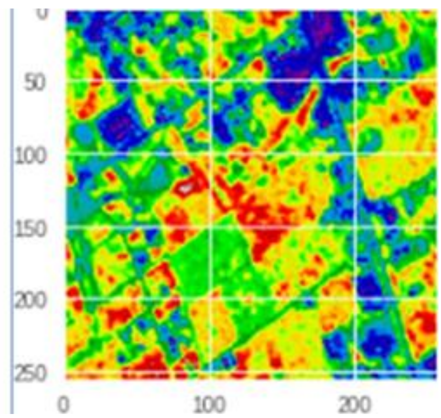

Fig-4b: Resultant Image in 2016

Fig-4a shows the original MODIS Satellite image that was captured in the year 2016 whereas Fig-4b shows the original MODIS Satellite image in 2016 after the classification i.e., resultant image.

Below test results are occurred by implementing TW-Dynamic Time Warping Algorithm.

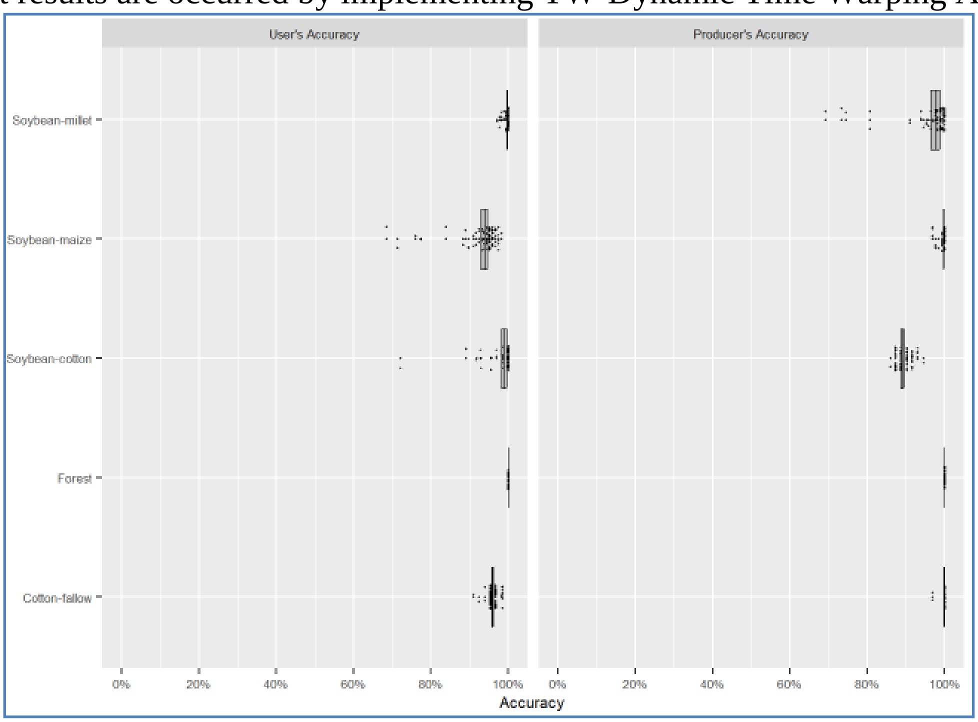

Fig-5: Cross Validation of User Accuracy and Producer Accuracy

Overall

Accuracy	sd	CImin	CImax
0.978	0.017	0.974	0.981

User's

	Accuracy	sd	CImin	CImax
Cotton-fallow	0.96	0.0119	0.96	0.96
Forest	1.00	0.0000	1.00	1.00
Soybean-cotton	0.99	0.0328	0.98	1.00
Soybean-maize	0.94	0.0476	0.93	0.95
Soybean-millet	1.00	0.0056	1.00	1.00

Producer's

	Accuracy	sd	CImin	CImax
Cotton-fallow	1.00	0.0046	1.00	1.00
Forest	1.00	0.0000	1.00	1.00
Soybean-cotton	0.89	0.0185	0.89	0.89
Soybean-maize	1.00	0.0074	1.00	1.00
Soybean-millet	0.98	0.0544	0.97	0.99

Fig-6: Accuracy Table

Fig-6 tells the overall accuracy obtained after implementing the TW-DTW algorithm i.e., 97.8% and it also describes the accuracy of each class label.

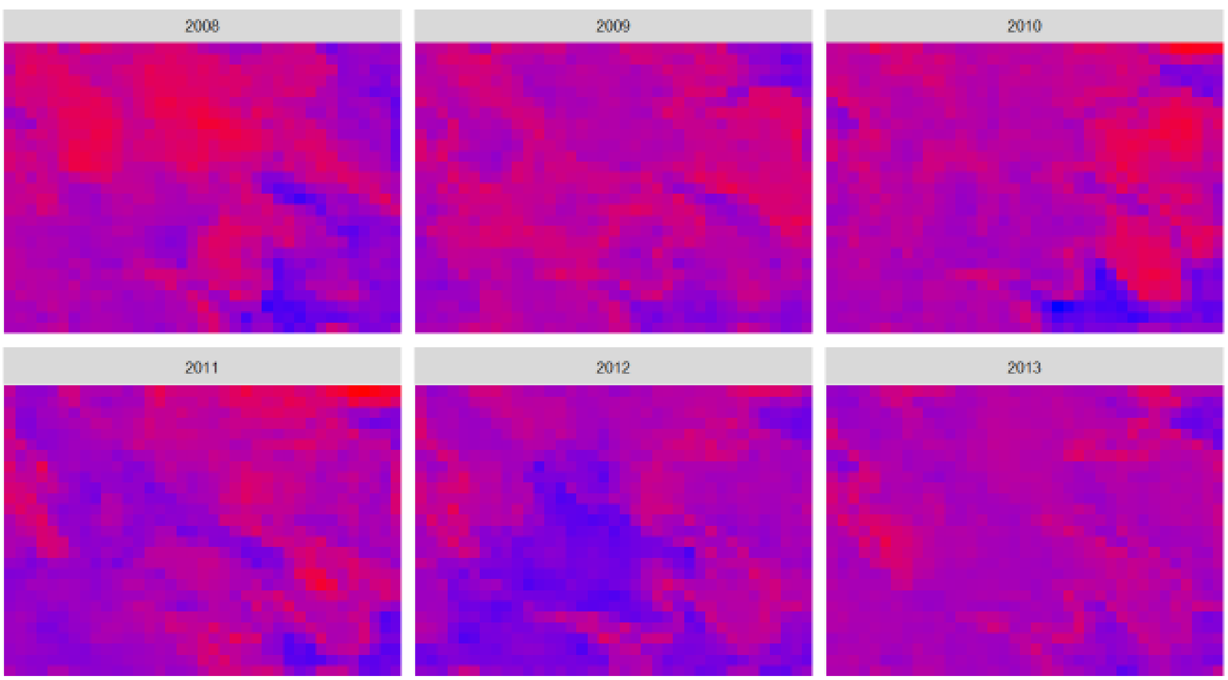

Fig-7: Change Map between 2 Time Stamps (2008-2013)

Fig-7 shows the change map obtained after the implementation TW-DTW, from 2008 to 2013. Here color gradients are used because it helps us in identifying the changes that occur annually.

8. Analysis of Performance:

The experiment findings demonstrate that CNN algorithm achieved an accuracy of 81.5% while the TW-DTW algorithm achieved accuracy 97.8% by combining the Linear Model and Logistic Model that results in formation of a hybrid model.

Table-3: Comparison between CNN and TWDTW

Algorithm	Accuracy (%)
CNN	81.5
TW-DTW	97.8

9. Result Analysis:

Promising results were obtained on integrating 2 methodologies together. The TW-DTW algorithm had a general correctness of 97.8% in comparison to CNN algorithm's 81.5%. While the TW-DTW method performed better at differentiating between agricultural and grassland areas, the CNN algorithm did better at differentiating between urban and forest areas. Overall, utilizing both of these methods instead of just one increased the conceptual framework is an analytical of land mapping. These findings imply that this strategy can be speedy and precise over very large swathes.

10. Conclusion and Future Study:

The strategies of both sets of algorithms for quantifying remotely sensed data utilize have yielded encouraging results. While the TWDTW algorithm has been successful in handling temporal and spectral fluctuations over time, the CNN algorithm has demonstrated excellent accuracy in identifying requisite resources from geospatial technology. In locations with complex and dynamic landscapes, these two techniques combined have the ability to boost the accurateness of mapping. However, more inquiry is necessary to confirm the proposed method's accuracy and its potential for widespread use.

The merging of both the algorithms for enhanced accuracy and efficiency could be used in future study for mapping of grazing lands and lithology. The desire to identify consumption of acreage throughout an area of interest, the study would include gathering high-resolution satellite imagery in addition with ground truth data. The resulting classified map would next be subjected to the TWDTW algorithm in order so as to consider temporal changes, which might be crucial in areas experiencing rapid development or natural disasters. The generated map would be contrasted with ground truth information and other regional maps in a bid to decide method. The study could potentially look at the method's effectiveness in various areas or in various temporal and environmental contexts. In general, this paper writing may have significant ramifications for land usage planning, natural resource management, and environmental monitoring by offering a more precise and effective method of mapping and tracking changes over time.

References:

[1] "Land-cover change identification using multi-temporal MODIS NDVI data," Remote Sens. Environ., vol. 105, no. 2, pp. 142–154, 2006; R. S. Lunetta, J. F. Knight, J. Ediriwickrema, J. G. Lyon, and L. D. Worthy.

[2] Identifying trend and seasonal variations in satellite image time series, J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, Remote Sens.Environ., vol. 114, no. 1, pp. 106-115, 2010.

[3] A pixel-based Landsat compositing approach for large area land cover mapping was published in the October 2013 issue of the IEEE Journal of Special Topics in Applied Earth Observation and Remote Sensing. The authors are P. Griffiths, S. van der Linden, T. Kuemmerle, and P. Hostert.

[4] Study of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central Great Plains, Remote Sens. Environ., vol. 108, no. 3, pp. 290-310, 2007. B. D. Wardlow, S. L. Egbert, and J. H. Kastens.

[5] Satellite image time series analysis under temporal warping, IEEE Trans. Geosci. Remote Sens., vol. 50, no. 8, pp. 3081–3095, Aug. 2012; F. Petitjean, J. Inglada, and P. Gancarski.

[6] Utilizing dynamic time warping to discover patterns in time series, D. J. Berndt and J. Clifford, Proc. KDD Workshop, 1994, pp. 359–370.

[7] Effective satellite image time series analysis under temporal warping, F. Petitjean and J. Weber, IEEE Geosci. Remote Sens. Lett., vol. 11, no. 6, pp. 114 3-11 47, June 2014.

[8] "Spatial statistical study of land-use determinants in the Brazilian Amazonia: Investigating intra-regional heterogeneity," Ecol. Model., vol. 209, nos. 2-4, pp. 169–188, 2007. A. P. D. Aguiar, G. Câmara, and M. I. S. Escada.

[9] Adam, E., Mutanga, O., Odindi, and Abdel-Rahman, E.M., 2014. Using RapidEye imagery, classifier performance of random forest and support vector machine is assessed in a varied coastal area. International Journal of Remote Sensing, 35(10), pp. 3440–3458.

[10] J.M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot (2013). Data analysis from hyperspectral remote sensing and upcoming difficulties. IEEE Geoscience and remote sensing magazine, 1(2), pp. 3–36.

[11] Alberti, M., Folke, C., Moran, E., Liu, J., Dietz, T., Carpenter, S. R., Taylor, W. W., et al (2007). linked human and natural systems' complexity. 317(5844):1513–1516 in Science.

[12] (Eds.) Lambin, E. F., and Geist, H. J. (2008). Changes in land use and land cover: local dynamics, global effects. Science & Business Springer.

[13] Turubanova, S. A., Tyukavina, A., Moore, R., Hancher, M., Hansen, M. C., and Townshend, J. R. (2013). worldwide maps illustrating the change in forest cover in the twenty-first century. 850–853 in Science, 342(6160).

[14] Brondizio, E., Mausel, P., and Moran, E. (2004). ways for spotting changes. 25(12), 2365–2407, International Journal of Remote Sensing.

[15] (Eds.) Seto, K. C., and Fragkias (2005). Changes in urban land use: a combination of socioeconomic and remote sensing data. Science & Business Springer.