iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.83228

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Satellite-based Change Detection and Analysis for Early Deforestation Prediction

Nilesh Sable, Anisha Gondhalekar, Sumiran Kadu, Deep Mehta, Lokesh Bhat
Computer Science and Engineering (AI) Vishwakarma Institute of Technology Pune, India

Abstract: Deforestation is a horrific environmental challenge which has direct effects on climate change, biodiversity and sustainable development. Deforestation can be identified by satellite imagery and interventions to be taken to prevent it are made in time when policy measures can be implemented. The present paper introduces a full satellite-based change detection and analysis framework of objectively predicting the deforestation at very early stages based on multi-temporal remote sensing measurements and the power of deep learning. The suggested system combines radiometric correction, analysis of vegetation index, and semantic segmentation based on the convolutional neural network in order to detect changes in the forest cover at an early phase. Multi-date satellite images are used to develop a temporal change detection pipeline that captures vegetation loss pattern variations that are subtle. Experimental findings using publicly available satellite datasets indicate that the proposed method experiences a high detection accuracy and strength relative to the conventional and classical machine learning methods based on thresholds. The effectiveness of the system in terms of early deforestation monitoring is confirmed by quantitative evaluation with the use of metrics like Intersection over Union (IoU), F1-score, and overall accuracy with the accuracy of 93.6 and IoU of 0.85.

Keywords: Deforestation Detection, Satellite Imagery, Change Detection, Deep Learning, Remote Sensing, CNN, Semantic Segmentation

I. INTRODUCTION

Anthropogenic processes such as urbanization, agricultural development, and unlawful logging activities have led to a geometric increase in global deforestation with disastrous ecological and climatic effects. According to the Food and Agriculture Organization (FAO), the annual loss of forests spreads over about 10 million hectares of forest, which is causing the rise of the greenhouse gases as well as loss of species and the interference with the hydrological processes.

The traditional forest monitoring methods are based on the periodic survey of a ground or manual observation of the satellites, which is not continuous and is costly and open to the bias of human interpretation. As the quality of satellite images has improved to produce high-resolution images, including Landsat, Sentinel, and commercial satellites, automated systems of deforestation detection are receiving growing research interest. Change detection by satellite is used to detect variations in land-cover by examining multi-temporal images of the same geographical location. Nevertheless, there are specific technical issues with early deforestation prediction due to the fact that early vegetation loss is revealed in the form of low spectral differences that are usually disrupted by seasonal phenological variations, atmospheric noise, different illumination geometry, and sensor calibration errors.

Recent developments in deep learning and in specific convolutional neural networks (CNNs) and attention-based systems have shown unprecedented accuracy in eliciting hierarchical spatial and temporal features of satellite imagery. The study will suggest a satellite-based, end-to-end system of monitoring vegetation change synergetically combining vegetation indices with a deep learning-based segmentation to identify deforestation at its initial stage. The algorithm is particularly suited to acquire the discriminative characteristics based on multi-temporal satellite information whilst preserving the strength in the case of natural seasonal changes and atmospheric perturbations.

II. LITERATURE REVIEW

Various methodological paradigms have been put forward towards detection of deforestation by the use of remote sensing data. Certain methods used in traditional approaches include the change detection of the vegetation indices like the Normalized Difference Vegetation Index (NDVI) with the use of threshold-based methods. The change detecting developed by Coppin et al. is pixel-based, implying that image differencing and ratioing technologies are employed, and it is highly sensitive to radiometric noise and changes in illumination.

The algorithms based on machine learning, such as Support Vector Machines (SVM) and Random Forests (RF) have been applied to categorize land-cover (change) but they require manually designed features and show little ability to learn more complex spatial patterns. More recent studies have examined deep learning models such as the Siamese CNN architecture of multi-temporal image comparison and Recurrent Neural Networks (RNNs) of modeling temporal relationships. However, the majority of the current systems are concerned with the detection of deforestation after it has occurred, but not predictive, early-stage detection, and do not usually extrapolate to a wide range of forest types and seasonal cycles.

III. OBJECTIVES

The main research problem of the given work consists in the creation of an efficient, automatic system of deforestation prediction in its initial stages, by means of multi-temporal satellite imaging. In particular, the problem is to identify real forest degradation and natural seasonal changes, atmospheric artifacts, and noise of sensors. The research goals are formulated as following:

1. Develop a preprocessing pipeline that can be successfully used to reduce radiometric variability and atmospheric disturbances during multi-temporal satellite collections.
2. Devise a composite framework that combines vegetation index analysis and a semantic segmentation approach based on deep learning that will be more successful in feature discrimination.
3. Train a CNN based-based architecture which can learn multi-date imagery with hierarchical spatial-temporal representations.
4. Carry out a detailed quantitative assessment and comparative study with classical and machine learning baseline techniques.

IV. NOVELTY AND CONTRIBUTIONS

The main originality of the study is the creation of a single, end-to-end system that integrates domain-specific environmental knowledge with the latest deep learning systems to predict deforestation in its initial stages. Compared to the current methods that consider vegetation analysis and change detection as separate modules, the proposed system adopts a closely integrated framework in which vegetation indices act as the attention systems to provide directions to the learning process of the neural network in terms of feature learning. The contributions of this work are specific and have been listed as as follows:

A. Hybrid Feature Extraction Architecture

The framework suggests a new hybrid feature extraction scheme, a combination of spectral vegetation indices (NDVI, EVI, SAVI) and convolutional features that have been learned using a channel-wise attention model. Conventional methods employ vegetation indices to preprocess or post-process classifications, and hence, their application in conjunction with deep learning models is restricted. Conversely, the proposed architecture takes into account vegetation indices as extra input channels to the CNN, which allows the network to acquire complementary representations that integrate handcrafted ecological features with data-driven spatial trends. The attention system adapts weighting of vegetation channels according to their discriminative information on the detection of deforestation, virtually ignoring noise due to seasonal phenological variation, and enhancing information present of anthropogenic forest degradation.

B. Temporal Normalization Framework

One of the technical contributions is a multi-scale temporal normalization framework that overcomes the underlying difficulty of differentiating actual loss of forests and the natural seasonal changes. Simple phenological cycles, atmospheric variations and sensor calibration drift are very prone to the conventional change detection methods requiring simple image differencing or ratioing. The temporally normalization scheme proposed is conforming to three levels of hierarchy (i) the normalization of pixel-radiometric is proposed based on radiometric correction of the pseudo-invariant feature (PIFs); (ii) the normalization of the histogram region is proposed; (iii) the normalization of the scene is proposed using Dark Object Subtraction (DOS) method. This multi-scale method results in a large decrease in false positive detections due to natural vegetation dynamics and maintains sensitivity to the anthropogenic disturbances of a subtle nature.

C. Early Prediction Capability Through Subtle Change Modeling

The system proposed is able to predict deforestation at early stages because it directly models minor temporal changes that precede major incidences of forest clearing. The majority of the current systems are proficient in detecting large-scale changes in land-cover in the post-facto, and, as a result, they are ineffective at detecting incipient deforestation, where selective logging, edge degradation, and partial canopy thinning occur.

Proposed CNN architecture includes a multi-resolution feature pyramid that contains change patterns on various scales of space allowing to detect localized removal of trees and slow forest fragmentation. Moreover, the training scheme is a temporally progressive augmentation scheme that progressively introduces the model to progressively finer change patterns, and thus makes the model more sensitive to early-stage degradation signals.

D. Robust Loss Function Design

As a means to overcome the class imbalance issue inherent in deforestation detection due to the large number of unchanged forest pixels compared to deforested ones, a new hybrid loss function is created. The loss function is a combination of pixel-wise classification with Binary Cross-Entropy (BCE) and region-based segmentation quality with Dice loss with weights composed of a spatially-adaptive term that focuses on hard-to-classify regions at the boundary.

This can be mathematically put as:

$L\_total = \alpha \cdot L\_BCE + \beta \cdot (1 - L\_Dice) + \gamma \cdot L\_boundary$

In which case the loss term of the boundary Lboundary is applied to misclassification in the boundary regions between forest and non-forest regions in particular, which are essential to detecting early deforestation. A dynamic adjustment of hyperparameters a, b and g is carried out during training according to the distribution of classes and the performance of detection, which is an adaptive curriculum learning strategy.

E. Comprehensive Comparative Analysis

Compared to the previous research that offers only a small amount of comparative analysis, the study in question is a thorough quantitative analysis of three methodological paradigms: (i) traditional threshold-based feature (NDVI thresholding, image differencing); (ii) classical machine learning (SVM, Random Forest); and (iii) deep learning (U-Net, Siamese CNN and the proposed hybrid model). The assessment includes several performance measures such as detection accuracy, computational efficiency, ability to perform generalization of various forest types, and resistance to atmospheric and seasonal changes. The statistical significance test with the help of McNemar test and confidence interval analysis can be taken as the strict validation of the superiority of the offered approach.

F. Practical Deployment Considerations

The presented system is developed keeping practical implementation limits in mind and includes model compression algorithms and effective inference channels that are applicable to practical forest monitoring systems. The architecture uses depthwise separable convolutions and knowledge distillation to minimize the computational complexity and can still achieve detection accuracy. Moreover, the system enables ability to learn in an incremental manner, so that the model can be updated periodically, with new satellite acquisitions and does not involve full retraining, thus, providing flexibility to the changing patterns of deforestation and new forms of forest disturbances.

V. SYSTEM ARCHITECTURE

The proposed change detection system is based on satellites and it consists of five major functional modules which are arranged in a hierarchical pipeline architecture. The system is constructed based on the principles of software engineering which is modular, and as such, is designed to be scalable, maintainable and can be extended to integrate with existing operational forest monitoring infrastructure. The architectural architecture focuses on the optimization of data flow, computational performance and capable error handling mechanisms. The overall system design is shown in a conceptual way and explained in detail in the following subsections. The satellite data acquisition module will handle the processing of incoming satellite signals to identify the position of the vehicle at any time.

A. Satellite Data Acquisition Module

The satellite data acquisition module will deal with the processing of the incoming satellite signals so that the position of the vehicle at any point in time will be known. David Data acquisition module connects to various satellite data archives such as the USGS Earth Explorer, ESA Copernicus Hub, and cloud-based systems such as Google Earth Engine. The module uses automated query functions on the basis of spatial region of interest (ROI), temporal range, cloud cover limitations and spectral band needs.

In this study, the multi-temporal information is obtained with the help of Landsat 8 OLI (Operation Land Imager) and Sentinel-2 MSI (MultiSpectral Instrument) sensors, with which, the data has complementary time resolution (16-day and 5-day revisit periods, respectively) and the spatial resolution (30m and 10m, respectively), respectively. Scenes are automatically selected by the module in quality terms such as the degree of cloud cover, completeness of data and sensor calibration. Metadata management subsystem has a complete database of the acquisition parameters such as sun elevation angle, sensor viewing geometry, atmospheric conditions and level of processing. Data ingestion pipeline supports Level-1 (top-of-atmosphere reflectance) and Level-2 (surface reflectance) products, and the automatic conversion and normalization of the data to a common data structure. The module produces multi-band raster datasets georeferenced and arranged in a time order, and the corresponding quality assurance (QA) bands to cloud masking and data validity testing.

B. Preprocessing and Radiometric Correction Module

The preprocessing module applies a full set of radiometric and geometric corrections that maintain a consistent multidate coverage and spatial orientation of multi-date imagery. The radiometric correction pipeline is made up of four consecutive processing steps:

1) Geometric Co-registration: Multi-temporal images are geometrically co-registered against a master reference image by automated feature-based co-registration. The module uses SIFT (Scale-Invariant Feature Transform) keypoint detection, RANSAC (Random Sample Consensus) outlier rejection as an approximation of an affine transformation matrix. Registration accuracy is checked, and the accuracy of the alignment is to the sub-pixel level of accuracy (RMSE less than 0.5 pixels).
2) Cloud and Shadow Masking: A cloud detection algorithm based on an ensemble cloud detection algorithm, Fmask (Function of mask) method with deep learning-based cloud classification is used to produce binary cloud masks. The shadow detection depends on the sun-sensor geometry and spectral analysis of dark pixels, which is based on the geometry of geometric modeling. QA layer identifies and removes detected clouds and shadows during the QA stage.
3) Atmospheric Correction: Level-1 products are atmospherically corrected using Dark Object Subtraction (DOS) method with a radiative transfer model based on MODTRAN. The correction takes the form of aerosol effects, Rayleigh scattering and water vapor absorptions, and transforms the top of the atmosphere reflectance to values of surface reflectance.
4) Relative Radiometric Normalization: Relative norm is used in order to address the issue of residual radiometric variation across acquisition dates by normalizing pseudo-invariant features (PIFs) like water bodies, bare soils, and man-made structures. PIFs between reference and target images are then fitted to a linear regression model and normalization coefficients obtained are used on a scene-wide basis.The mathematical formulation of radiometric normalization is expressed as:

$$\rho'\_\lambda = (\rho\_\lambda - \mu\_\lambda) / \sigma\_\lambda$$

where ρ_λ represents the raw reflectance in spectral band λ, μ_λ and σ_λ denote the mean and standard deviation computed from pseudo-invariant features, and ρ'_λ is the normalized reflectance. This z-score normalization ensures zero-mean, unit-variance distributions across temporal acquisitions, facilitating robust change detection.

C. Vegetation Index Computation Module

Computation module of vegetation index is used to compute various spectral indices that are intended to maximize vegetation features and minimize non-vegetation signals. Three indices of vegetation complementary to each other are calculated:

Normalized Difference Vegetation Index (NDVI): NDVI takes advantage of the high relative to near-infrared (NIR) reflectance (high when vegetation is healthy, low when chlorophyll absorbs light) and red reflectance (low when vegetation is healthy, high when chlorophyll absorbs light). The formulation is:

$$NDVI = (\rho\_NIR - \rho\_RED) / (\rho\_NIR + \rho\_RED)$$

The values of NDVI vary between -1 and +1 where the values of over 0.6 are dense vegetation, values between 0.2 and 0.6 are sparse vegetation, and values less than 0.2 represent bare soil or non-vegetated surfaces.

Enhanced Vegetation Index (EVI): EVI uses blue band reflectance to remove aerosol scattering, and is not as saturated in high-biomass areas:

$$EVI = G \times [(\rho\_NIR - \rho\_RED) / (\rho\_NIR + C1 \times \rho\_RED - C2 \times \rho\_BLUE + L)]$$

where G = 2.5 (gain factor), C1 = 6, C2 = 7.5 (aerosol resistance coefficients), and L = 1 (canopy background adjustment).

Soil-Adjusted Vegetation Index (SAVI): SAVI minimizes soil brightness influences in areas with sparse vegetation:

$$SAVI = [(\rho\_NIR - \rho\_RED) / (\rho\_NIR + \rho\_RED + L)] \times (1 + L).$$

where L = 0.5 of average vegetation cover. The module determines the change in vegetation indices (ΔNDVI, ΔEVI, ΔSAVI) between the successive dates of acquisition and it produces maps of change magnitude used to quantify the amount of vegetation change. These indices are input features into the deep learning module that give ecologically relevant representations that supplement raw spectral bands.

D. Deep Learning-based Change Detection Module

The deep learning-based change detection module is the core analytical part of the system and it consists of modified U-Net architecture, including attention mechanisms and multi-scale feature fusion. These construction principles are driven by the needs of: (i) pixel-accurate segmentation of deforested areas; (ii) the ability to incorporate multi-temporal data; (iii) the ability to capture both local texture structures and global contextual cues; and (iv) computational efficiency to be used in large scale applications.

1) Network Architecture: The network architecture contains three main components: an encoder pathway containing hierarchical feature extraction, a decoder pathway containing spatial resolution recovery and skip connections containing multi-scale feature fusion. The encoder has four convolutional blocks, each block comprises of two 3x3 convolutional blocks with two batch normalization blocks, ReLU activation and two 2x2 max-pooling blocks to downsample the spatial location. The receptive field successively increases in the size of 3x3 pizepels in the first layer to 131x131 pizepels in the bottleneck layer, allowing the capture of fine-grained texture as well as massive spatial context.
2) Temporal Feature Integration: Multi-temporal images are run through a Siamese encoder model that shares weights to produce individual feature representations of each acquisition date. Early concatenation of multi-date features is also done at every encoder stage followed by 1x1 convolutional layers to reduce dimensionalities and refine features. In this design, the network is allowed to learn temporal pattern of correlation, but keep the spatial alignment.
3) Attention Mechanisms: Each level of the decoder has the channel attention modules, which selectively focus on the discriminative channels of features. The weights of attention are calculated using global average pooling and then a two-layer multi-layer perceptron (MLP), so that there can be adaptive recalibration of features according to the features that are important in change detection. During the initial phases of deforestation, vegetation index channels are given more attention weight, whereas the spectral band channels take the center stage in the later stages of massive forest clearance.
4) Output Generation: The decoder generates a probability map, $P(x,y) \in [0,1]$which represents the probability of the pixel being deforested. An activation function that is a sigmoid is used at the output layer in order to limit the predictions to the valid range of probabilities. Adaptive thresholding by using the approach of Otsu is used to produce the final binary change map, which automatically decides the best threshold by maximizing inter-class variance.

E. Post-processing and Visualization Module

The latter post-processing component completes the five main morphological operations with spatial filters refining changes in the entire raw change detection output. Where there are small isolated detections (< 0.5 hectares) that are probably classification noise, area-based filtering removes these small detections. The small areas adjacent to a detected deforestation patch are filled with morphological closing operators using circular structuring elements (radius = 3 pixels) to create spatially coherent areas of change. Connected component analysis appends tags to single instances of deforestation and calculates geometric measures such as area, perimeter, shape complexity and spatial distribution measures. The visualization subsystem produces multi-layer cartographic products such as false-color composite imagery, heatmap of change magnitude, temporal sequences of animation, interactive web based dashboards and so forth. Identified deforestation polygons are vectorized and available as standard GIS formats (Shapefile, GeoJSON, KML) to be integrated with forest management information system. Enhanced mechanisms of real timer triggering expect prioritized deforestation events, including data of the area covered, new neighborhoods and the rates of change in time, so that timely response may be undertaken by forest protection departments.

F. System Integration and Data Flow

The entire system is rolled out as a distributed processing chain with Apache Spark to process satellite imagery on a large scale. The architecture allows operating both batch processing to do historical analysis and near real time processing to monitor operations. Intermediate outputs of every processing step are put into a distributed file system (HDFS) so that they can be fault tolerant to development through iteration. To integrate with external access to monitoring forests, the system has been opened to external API calls which may also be systematically added via the Apache Airflow system to run periodic deals involving new satellite purchases.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Performance monitoring and logging systems trace processing throughput, detection latency, resource consumption associated with the system processes, along with this way the system meets a guarantee through operations that are reliable and constant processing resource optimization as it continues to operate.

VI. METHODOLOGY

The methodological framework includes the mathematical expressions, procedures of algorithms, and experimental plan that is utilized in the detection of change as well as early deforestation by satellites. This chapter gives a detailed technical information about theoretical basis, model structures, training systems and testing guidelines.

A. Problem Formulation and Mathematical Framework

Let $I = \{I_{t1}, I_{t2}, ..., I_{tn}\}$ represent a multi-temporal satellite image sequence of a geographical region, where $I_{tk} \in \mathbb{R}^{(H\times W\times B)}$ denotes a satellite image acquired at time t_k, with H×W spatial resolution and B spectral bands. Each pixel $p = (i,j)$ in the image is characterized by a spectral signature vector $s_p = [\rho_1, \rho_2, ..., \rho_B]$, where ρ_b represents the reflectance value in spectral band b.

The objective of change detection is to learn a mapping function $f: \mathbb{R}^{(H\times W\times B)} \times \mathbb{R}^{(H\times W\times B)} \rightarrow \{0,1\}^{(H\times W)}$ such that:

$$C = f(I_{ti}, I_{tj})$$

where $C \in \{0,1\}^{(H\times W)}$ represents a binary change map indicating deforestation ($C(i,j) = 1$) or no-change ($C(i,j) = 0$) between acquisition times t_i and t_j. The proposed framework decomposes f into two sequential transformations:

$$C = h(g(I_{ti}), g(I_{tj}))$$

where $g: \mathbb{R}^{(H\times W\times B)} \rightarrow \mathbb{R}^{(H\times W\times D)}$ represents the vegetation feature extraction function that augments raw spectral bands with computed vegetation indices ($D > B$), and $h: \mathbb{R}^{(H\times W\times 2D)} \rightarrow \{0,1\}^{(H\times W)}$ denotes the deep neural classifier that produces the binary change map.

B. Radiometric Preprocessing and Normalization

Radiometric preprocessing is meant to create radiometric consistency radiometric consistency between acquires done in multiple temporal periods by compensation of atmospheric factors, sensor calibration variation, and illumination variation. The multi-stage normalization procedure is used in the preprocessing pipeline.

1) Atmospheric Correction

Atmospheric correction Level-1 top-of-atmosphere (TOA) reflectance products are atmospheric products that can be atmospherically corrected through Dark Object Subtraction (DOS) method.The surface reflectance $\rho_{surface}$ is estimated from TOA reflectance ρ_{TOA} as:

$$\rho_{surface,\lambda} = (\rho_{TOA,\lambda} - \rho_{path,\lambda}) / T_{atm,\lambda}$$

where $\rho_{path,\lambda}$ represents the path radiance (estimated from dark object pixels, typically < 1% TOA reflectance), and $T_{atm,\lambda}$ denotes atmospheric transmittance (modeled using MODTRAN lookup tables based on sensor viewing geometry and atmospheric conditions).

2) Relative Radiometric Normalization

Relative normalization is used to establish radiometric correspondence between the subject image $I_{subject}$ and the reference image I_{ref} by means of linear regression of the pseudo-invariant features (PIFs).The transformation of normalization can be calculated:

$$\rho'_{subject,\lambda} = a_\lambda \cdot \rho_{subject,\lambda} + b_\lambda$$

where coefficients a_λ and b_λ are determined by ordinary least squares (OLS) regression using PIFs. PIFs are determined by an automated algorithm which picks up the stable pixels (coefficient of variation less than 10% among all acquisition dates) corresponding to different land cover types.

3) Z-score Standardization

Following atmospheric and relative corrections, z-score standardization is applied to ensure zero-mean, unit-variance distributions, facilitating convergence during neural network training:

$$\rho''_\lambda = (\rho'_\lambda - \mu_\lambda) / \sigma_\lambda$$

where μ_λ and σ_λ represent the mean and standard deviation of band λ computed from the entire image or from a representative sample of pixels.

C. Vegetation Index Computation and Feature Engineering

Normalized reflectance is then used to compute vegetation indices in order to improve spectral properties of vegetation. The three main indices used are mathematically stated as follows.

1) Normalized Difference Vegetation Index (NDVI)

The NDVI takes advantage of the chlorophyll absorption of red light and the widely reflective near infrared light at the relatively longer wavelengths produced and used by vegetation:

$$NDVI = (\rho\_NIR - \rho\_RED) / (\rho\_NIR + \rho\_RED + \varepsilon)$$

where ε = 10^-8 is a regularization term preventing division by zero. NDVI ranges from -1 (water, clouds) to +1 (dense vegetation), with typical forest values exceeding 0.6. Temporal NDVI differences ΔNDVI = NDVI_t2 - NDVI_t1 quantify vegetation change magnitude, with negative values indicating potential deforestation.

2) Enhanced Vegetation Index (EVI)

EVI uses blue band reflectance to decorrect atmospheric aerosol scattering and has lower saturation in biomass-rich regions:

$EVI = 2.5 \times [(\rho\_NIR - \rho\_RED) / (\rho\_NIR + 6\rho\_RED - 7.5\rho\_BLUE + 1)]$

EVI is more sensitive to changes in structure in dense canopy forest than NDVI, and therefore find application specifically in tropical forest monitoring environments where biomass saturation is a major concern.

3) Soil-Adjusted Vegetation Index (SAVI)

SAVI minimizes soil background influences in sparsely vegetated regions through an adjustment factor L:

$SAVI = [(\rho\_NIR - \rho\_RED) / (\rho\_NIR + \rho\_RED + L)] \times (1 + L)$

where and L = 0.5 is the average vegetative density assumption. SAVI is especially useful in tracking intimate deforestation in the forest-savanna transition areas with the increasing soil exposure.

D. Classical Machine Learning Baseline Methods

In terms of offering a full comparative analysis, two traditional machine learning algorithms are used as baseline algorithms: Support Vector Machines (SVM) and Rand Forest (RF) classifiers.

1) Support Vector Machine (SVM)

The goal of SVM is to identify a separating hyperplane in feature space which maximizes the distance between the forest and non-forest categories.Given labeled training samples {(x_i, y_i)}^N_i=1 where x_i∈ℝ^d represents the feature vector (concatenation of spectral bands and vegetation indices) and y_i∈ {-1, +1} denotes the class label, the SVM optimization problem is formulated as:

$$min\_(w,b,\xi)\ (1/2)||w||^2 + C\ \Sigma^N\_i = 1\ \xi\_i$$

subject to the constraints: y_i(w·x_i + b) ≥ 1 - ξ_i and ξ_i ≥ 0, In which w is the weight vector perpendicular to the hyperplane, b is a bias factor, xi are slack variables which allow non-separable data to be custom fitted, and C is the regularization parameter which determines the compromise between maximizing the margin and minimising training errors.A radial basis function (RBF) kernel K(x_i, x_j) = exp(-γ||x_i - x_j||^2) is employed to enable nonlinear decision boundaries, with kernel parameter γ = 1/(2σ^2) determining the influence radius. Hyperparameters C and γ are optimized through 5-fold cross-validation grid search over logarithmically spaced ranges: C ∈ {10^-2, 10^-1, 10^0, 10^1, 10^2} and γ ∈ {10^-4, 10^-3, 10^-2, 10^-1, 10^0}.

2) Random Forest (RF)

Random Forest builds a complex of bootstrap samples of the original data decision trees{T_k}^K_k=1 trained on bootstrap samples of the original dataset. Each tree T_k is grown by recursively partitioning feature space based on information gain criteria. At each split node, A random sample of m features (m = √d in classification problems, d being the overall feature dimensionality) is also tested at every split node and the feature with the highest information gain is chosen to partition.IG (n) of a node n uses information gain to compute a split on feature f at that node.

$$IG(n,f) = H(n) - \Sigma\_(c \in \{left, right\})\ (|n\_c|/|n|) \cdot H(n\_c)$$

where we have H(n) is the Gini impurity or entropy of node n, n_c is child nodes of a split. Growth of the tree proceeds till leaf nodes meet a stopping criterion (minimal sample size, maximum depth or a purity level). The ultimate RF prediction is arrived at via majority prediction amongst all the trees:

$$\hat{y}(x) = mode\{T\_1(x), T\_2(x), \ldots, T\_K(x)\}$$

Cross-validation is used to optimize the ensemble size K and maximum tree depth, and common values of K = 100 trees and depth = 20 levels get optimal biasvariance trade off in the deforestation detection problem.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

E. Convolutional Neural Network Architecture

The proposed deep learning architecture is based on a modified U-Net with channel attention mechanisms and multi-scale feature fusion. The architecture comprises an encoder-decoder structure with skip connections facilitating information flow between corresponding resolution levels.

1) Convolutional Layer Operations

A convolutional layer applies learnable filters to extract local spatial patterns from input feature maps. For an input feature map $X \in \mathbb{R}^{H\times W\times C_{in}}$ and a convolutional kernel $K \in \mathbb{R}^{k\times k\times C_{in}\times C_{out}}$, the output feature map $Y \in \mathbb{R}^{H'\times W'\times C_{out}}$ is computed as:

$$Y(i,j,c) = \sum_{c'=1}^{C_{in}} \sum_{m=1}^{k} \sum_{n=1}^{k} X(i+m-\lceil k/2\rceil, j+n-\lceil k/2\rceil, c') \cdot K(m,n,c',c) + b_c$$

where b_c represents the bias term for output channel c, and $\lceil\cdot\rceil$ denotes the ceiling function. Zero-padding is applied to maintain spatial dimensions. The network employs 3×3 kernels throughout, balancing receptive field size with computational efficiency.

2) Activation Functions

Rectified Linear Unit (ReLU) activation is applied after each convolutional layer to introduce nonlinearity:

$$ReLU(x) = max(0, x)$$

ReLU addresses the vanishing gradient problem and accelerates convergence compared to sigmoid or tanh activations. The output layer employs a sigmoid activation to produce pixel-wise probabilities:

$$\sigma(z) = 1 / (1 + e^{-z})$$

where z represents the pre-activation logit value, and $\sigma(z) \in [0,1]$ indicates the deforestation probability.

3) Pooling Operations

Max-pooling with 2×2 windows and stride 2 is employed for spatial downsampling in the encoder pathway:

$$Y(i,j,c) = max\{X(2i+m, 2j+n, c) : m,n \in \{0,1\}\}$$

Max-pooling provides translation invariance and reduces computational complexity while preserving salient features. The decoder pathway employs transposed convolutions (deconvolutions) for spatial upsampling, learning optimal interpolation kernels during training.

4) Batch Normalization

Batch normalization is applied after each convolutional layer to stabilize training and accelerate convergence. For a mini-batch of feature map activations x_i, the normalized value $\hat{x}_i$ is computed as:

$$\hat{x}_i = (x_i - \mu_B) / \sqrt{(\sigma^2_B + \varepsilon)}$$

followed by an affine transformation $y_i = \gamma\hat{x}_i + \beta$, where μ_B and σ^2_B represent batch mean and variance, $\varepsilon = 10^{-5}$ is a numerical stability constant, and γ, β are learnable parameters. Batch normalization reduces internal covariate shift and enables higher learning rates.

5) Channel Attention Mechanism

Channel attention modules adaptively recalibrate feature channels based on their discriminative importance. For input feature map $F \in \mathbb{R}^{H\times W\times C}$, the channel attention weights $\alpha \in \mathbb{R}^C$ are computed through:

$$\alpha = \sigma(W_2 \cdot \delta(W_1 \cdot GAP(F)))$$

where $GAP(\cdot)$ represents global average pooling that aggregates spatial information, $W_1 \in \mathbb{R}^{C/r\times C}$ and $W_2 \in \mathbb{R}^{C\times C/r}$ are weight matrices with reduction ratio $r = 16$, δ is ReLU activation, and σ is sigmoid activation. The attention-modulated feature map F' is obtained as:

$$F'(i,j,c) = \alpha_c \cdot F(i,j,c)$$

This mechanism enables the network to emphasize vegetation-related channels during early deforestation detection while suppressing irrelevant or noisy channels.

6) Network Depth and Capacity

The encoder consists of four encoding blocks with feature channel dimensions [32, 64, 128, 256], progressively capturing multi-scale representations.

The bottleneck layer has 512 channels, providing a compressed latent representation. The decoder mirrors the encoder structure with corresponding channel dimensions [256, 128, 64, 32]. Skip connections concatenate encoder and decoder features at matching resolution levels, preserving fine-grained spatial details lost during downsampling. The total network contains approximately 8.2 million trainable parameters.

F. Loss Function Formulation

The training objective combines multiple complementary loss components to address class imbalance and optimize both pixel-wise accuracy and region-based segmentation quality.

1) Binary Cross-Entropy Loss

Binary cross-entropy (BCE) measures pixel-wise classification error:

$$L\_BCE = -(1/N)\, \Sigma^N\_(i = 1)\, [y\_i \log(p\_i) + (1 - y\_i) \log(1 - p\_i)]$$

where $N = H\times W$ represents total pixel count, $y\_i \in \{0,1\}$ is the ground truth label, and $p\_i \in [0,1]$ is the predicted probability for pixel i. To address class imbalance (deforested pixels typically constitute < 5% of total pixels), weighted BCE is employed:

$$L\_BCE\_weighted = -(1/N)\, \Sigma^N\_(i = 1)\, [w\_pos \cdot y\_i \log(p\_i) + w\_neg \cdot (1 - y\_i) \log(1 - p\_i)]$$

where $w\_pos = N\_neg/N\_total$ and $w\_neg = N\_pos/N\_total$ represent inverse class frequency weights, ensuring balanced contribution from minority class.

2) Dice Loss

Dice loss optimizes region-based overlap between predicted and ground truth segmentation masks:

$$L\_Dice = 1 - (2|P \cap G| + \varepsilon) / (|P| + |G| + \varepsilon)$$

where P represents the predicted segmentation set, G is the ground truth set, $|\cdot|$ denotes cardinality (sum of probabilities), and $\varepsilon = 1$ is a smoothing term preventing division by zero. Dice loss is particularly effective for imbalanced segmentation tasks as it directly optimizes the Dice coefficient (F1-score), emphasizing correct detection of positive class regions.

3) Boundary Loss

A boundary loss term emphasizes accurate delineation of deforestation boundaries, which are critical for early detection. The boundary loss is computed as:

$$L\_boundary = -(1/N\_b)\, \Sigma\_(i \in boundary)\, [y\_i \log(p\_i) + (1 - y\_i) \log(1 - p\_i)]$$

where N_b represents the number of boundary pixels (identified by morphological gradient operation on ground truth masks), and the summation is restricted to pixels in boundary regions. This formulation assigns higher penalty to misclassifications in transitional zones between forest and non-forest areas.

4) Combined Loss Function

The total loss function is a weighted combination of BCE, Dice, and boundary losses:

$L\_total = \alpha \cdot L\_BCE + \beta \cdot L\_Dice + \gamma \cdot L\_boundary$

where $\alpha = 0.4$, $\beta = 0.4$, and $\gamma = 0.2$ are empirically determined hyperparameters. This composite loss function balances pixel-wise classification accuracy (BCE), region-based segmentation quality (Dice), and boundary precision (boundary loss), yielding superior performance for early deforestation detection compared to individual loss components.

G. Training Procedure and Optimization

The CNN model is trained using the Adam (Adaptive Moment Estimation) optimizer, which combines momentum-based gradient descent with adaptive learning rates for individual parameters. The Adam update rule for parameter θ at iteration t is:

$$\theta\_t = \theta\_(t-1) - \eta \cdot \hat{m}\_t / (\sqrt{\hat{v}\_t} + \varepsilon)$$

where $\hat{m}\_t = m\_t/(1-\beta_1^t)$ and $\hat{v}\_t = v\_t/(1-\beta_2^t)$ represent bias-corrected first and second moment estimates, computed as exponential moving averages: $m\_t = \beta_1 \cdot m\_(t-1) + (1-\beta_1) \cdot g\_t$ and $v\_t = \beta_2 \cdot v\_(t-1) + (1-\beta_2) \cdot g\_t^2$, where g_t is the gradient at iteration t. Hyperparameters are set as: initial learning rate $\eta = 1\times10^{-4}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$.

The learning rate is scheduled using a cosine annealing strategy with warm restarts, promoting escape from local minima and enhancing convergence. The learning rate at epoch e is computed as:

$\eta\_e = \eta\_min + (\eta\_max - \eta\_min) \cdot (1 + \cos(\pi T\_cur/T\_max)) / 2$

where T_cur represents the current epoch within the restart cycle, T_max = 50 is the cycle length, $\eta\_max = 1\times10^{-4}$ and $\eta\_min = 1\times10^{-6}$. Training is conducted for 200 epochs with batch size 16, utilizing mixed precision (FP16) training to reduce memory consumption and accelerate computation on GPUs.

H. Data Augmentation

To enhance model generalization and robustness, extensive data augmentation is applied during training. Geometric transformations include random rotation (±90°), horizontal and vertical flipping (probability 0.5), and random cropping to 256×256 patches. Photometric augmentations include random brightness adjustment (±15%), contrast variation (0.8-1.2), and Gaussian noise injection ($\sigma = 0.01$). Importantly, temporal augmentations are employed where the order of image pairs is randomly swapped with probability 0.3, encouraging the model to learn symmetric temporal relationships rather than overfitting to specific acquisition sequences.

I. Evaluation Metrics

Model performance is quantitatively assessed using multiple complementary metrics that capture different aspects of detection quality.

Overall Accuracy: ACC = (TP + TN) / (TP + TN + FP + FN), where TP, TN, FP, FN represent true positives, true negatives, false positives, and false negatives, respectively.

Precision: Precision = TP / (TP + FP), measuring the proportion of correctly identified deforested pixels among all pixels classified as deforested.

Recall (Sensitivity): Recall = TP / (TP + FN), quantifying the proportion of actual deforested pixels correctly identified.

F1-score: F1 = 2·(Precision·Recall) / (Precision + Recall), providing the harmonic mean that balances precision and recall.

Intersection over Union (IoU): IoU = TP / (TP + FP + FN), measuring the overlap between predicted and ground truth deforestation regions. IoU is particularly informative for segmentation tasks as it simultaneously penalizes both false positives and false negatives.

J. Dataset Description and Experimental Setup

Experiments are conducted on multi-temporal Landsat 8 and Sentinel-2 imagery covering diverse tropical and temperate forest regions including the Amazon Basin (Brazil), Congo Basin (Central Africa), and Southeast Asian rainforests (Indonesia, Malaysia). The dataset comprises 15,000 image pairs with spatial resolution 10-30m, acquired between 2018-2024 with temporal baselines ranging from 1 month to 2 years. Ground truth deforestation labels are generated through manual digitization by expert interpreters, validated against field surveys and high-resolution commercial satellite imagery. The dataset is partitioned into training (70%, 10,500 pairs), validation (15%, 2,250 pairs), and testing (15%, 2,250 pairs) sets, stratified by geographic region to ensure balanced representation.

All experiments are executed on an NVIDIA A100 GPU (40GB memory) using PyTorch 2.0 framework. Model training requires approximately 18 hours for 200 epochs, while inference on a 1000×1000 pixel image requires 0.8 seconds, enabling near-real-time operational deployment.

VII. RESULTS AND DISCUSSION

This section presents comprehensive experimental results, comparative analysis, and critical discussion of the proposed early deforestation prediction system. The evaluation encompasses quantitative performance metrics, qualitative visual assessment, ablation studies, generalization analysis, computational efficiency, and failure case analysis.

A. Quantitative Performance Evaluation

Table 1 presents a comprehensive quantitative comparison of the proposed CNN-based method against baseline approaches including traditional threshold-based methods and classical machine learning techniques.

Table 1: Comparative Performance Analysis on Test Dataset

Method	Accuracy	Precision	Recall	F1-score	IoU
NDVI Thresholding	82.4	0.76	0.80	0.78	0.64

Image Differencing	79.8	0.72	0.78	0.75	0.60
Support Vector Machine	86.9	0.82	0.84	0.83	0.71
Random Forest	88.2	0.84	0.86	0.85	0.74
U-Net (Baseline)	91.2	0.88	0.87	0.88	0.78
Proposed Method	93.6	0.92	0.90	0.91	0.85

The results demonstrate that the proposed CNN-based method significantly outperforms all baseline approaches across all evaluation metrics. The achieved accuracy of 93.6% represents a 5.4 percentage point improvement over the best-performing classical machine learning method (Random Forest: 88.2%) and a 2.4 percentage point gain over standard U-Net architecture (91.2%). The superior IoU score of 0.85 indicates excellent spatial agreement between predicted and ground truth deforestation regions, confirming the system's capability for precise boundary delineation.

B. Ablation Study: Component Contribution Analysis

An ablation study was conducted to quantify the individual contributions of key architectural components. Table 2 presents performance metrics for different model configurations.

Table 2: Ablation Study Results

Model Configuration	Accuracy (%)	F1-score	IoU
Baseline (spectral bands only)	89.8	0.86	0.75
+ Vegetation indices	91.4	0.88	0.79
+ Channel attention	92.6	0.89	0.82
+ Hybrid loss function (Full model)	93.6	0.91	0.85

The ablation study reveals that each architectural component provides substantial performance gains. Incorporation of vegetation indices improves accuracy by 1.6 percentage points and IoU by 0.04, demonstrating the value of domain-specific ecological features. Channel attention mechanisms contribute an additional 1.2 percentage point accuracy improvement and 0.03 IoU gain, confirming that adaptive feature weighting enhances discriminative capacity. The hybrid loss function provides the final 1.0 percentage point accuracy boost and 0.03 IoU improvement, validating the importance of addressing class imbalance and boundary precision simultaneously.

C. Early-stage Deforestation Detection Capability

A critical evaluation dimension is the system's performance on early-stage deforestation events characterized by subtle vegetation changes. The test dataset is stratified into three temporal categories based on the extent of forest loss: early-stage (0-20% canopy loss), intermediate-stage (20-50% loss), and advanced-stage (>50% loss). Table 3 presents detection performance across these categories.

Table 3: Performance by Deforestation Stage

Deforestation Stage	Proposed Method F1	Random Forest F1	Improvement
Early-stage (0-20%)	0.83	0.71	+0.12
Intermediate (20-50%)	0.91	0.85	+0.06
Advanced (>50%)	0.96	0.93	+0.03

The results demonstrate that the proposed method exhibits particularly strong performance on early-stage deforestation detection, achieving F1-score of 0.83 compared to 0.71 for Random Forest, representing a 16.9% relative improvement. This substantial performance gap in early-stage detection validates the system's capability for predictive monitoring rather than merely post-facto assessment. The performance advantage diminishes for advanced-stage deforestation (3.2% relative improvement), where changes are sufficiently pronounced that even simpler methods achieve high accuracy. These findings confirm that the hybrid architecture with vegetation index integration and attention mechanisms is specifically effective for detecting subtle forest degradation patterns.

D. Generalization Analysis Across Forest Types

To assess model generalization capacity, performance is evaluated separately on three major forest ecosystems: tropical rainforests, temperate deciduous forests, and boreal coniferous forests. Each ecosystem presents distinct challenges including different canopy structures, seasonal patterns, and spectral characteristics.

Table 4: Cross-ecosystem Generalization Performance

Forest Ecosystem	Accuracy (%)	F1-score	IoU
Tropical Rainforest	94.2	0.92	0.86
Temperate Deciduous	92.8	0.89	0.83
Boreal Coniferous	91.4	0.88	0.81

The model demonstrates robust generalization across diverse forest ecosystems, with accuracy varying by only 2.8 percentage points (91.4%-94.2%) across biomes. The highest performance is achieved in tropical rainforests (94.2% accuracy), likely due to strong NDVI contrast between dense canopy vegetation and cleared areas, combined with minimal seasonal variations. Temperate deciduous forests exhibit slightly lower performance (92.8%), attributable to seasonal leaf-off periods that reduce vegetation index discriminability. Boreal forests present the greatest challenge (91.4%) due to lower vegetation density, extensive cloud cover, and snow interference. Nevertheless, the consistent high performance across all ecosystems validates the model's generalization capacity and suitability for global-scale deployment.

E. Computational Efficiency Analysis

Operational forest monitoring systems require efficient processing to enable timely intervention. Table 5 presents computational performance metrics for different methods.

Table 5: Computational Performance Comparison

Method	Training Time (hours)	Inference (1000×1000 px)
Random Forest	2.3	3.2s
U-Net (Baseline)	14.2	0.6s
Proposed Method	18.0	0.8s

The proposed method achieves inference time of 0.8 seconds per 1000×1000pixel image on an NVIDIA A100 GPU, enabling processing of approximately 4,500 km² of satellite imagery per hour (assuming 30m spatial resolution). While training requires 18 hours on the full dataset, this is a one-time cost, and the trained model can be deployed for operational monitoring. The inference speed is comparable to baseline U-Net (0.6s) despite additional architectural complexity, demonstrating that the performance gains do not come at prohibitive computational cost. For comparison, Random Forest requires 3.2s inference time, indicating that deep learning approaches provide superior accuracy-efficiency trade-offs for large-scale deployment.

F. Statistical Significance Testing

McNemar's test is used to evaluate the statistical significance of the difference between the proposed method and other standard methods. The null hypothesis H_0 assumes the error rates of the two methods are the same. Random Forest is used for comparison and the test statistic is $\chi^2 = 48.6$ ($p < 0.001$) which means that the enhancement of performance is statistically significant at the 99.9% confidence level. Likewise, a comparison with U-Net gives $\chi^2 = 22.4$ ($p < 0.001$) and shows significant superiority. These findings are statistically robust when one considers that the pattern of improvements in the observed performance is not random chance, but rather results from methodological improvements.

G. False Positive and False Negative Analysis

The qualitative analysis of misclassifications patterns shows unique properties of false positive and false negative patterns. The false positives (pixels incorrectly identified as deforested) mainly occur in the following cases: (i) agricultural crop rotation which causes a reduction in the amount of vegetation (32% of false positives); (ii) seasonal senescence in deciduous forests, a change of seasons (autumn to winter, or vice versa) (28% of false positives); (iii) natural disturbance, such as wildfire burn scars and storm damage (22% of false positives); and (iv) shadows cast by topographic relief or clouds (18% of false positives). These false positive findings indicate improvements to be made with the addition of multi-temporal context (permanent vs. temporary changes) and auxiliary terrain data.

The false negatives, or those incidents of actual deforestation which did not appear in the system, have different characteristics: (i) selective logging with partial retention of the forest canopy (42% of the false negatives); (ii) small-scale deforestation below the spatial resolution threshold (<0.1 hectares, 26%); (iii) gradual decrease in forest cover through removal of the understory with no change to the canopy cover (19%); (iv) deforestation covered by persistent cloud cover (13%). Selective logging is pervasive in false negatives, and this is the basic problem of detecting subtle forest degradation which is still a research frontier. These limitations could be overcome in future work by integrating LiDAR data with the high-resolution commercial imagery.

H. Comparison with State-of-the-art Literature

The results of the proposed system are compared with recently published methods in the literature. A global time series forest change detection based on Landsat data provided by Hansen et al. (2013) reported 65.9% user's accuracy (which is equivalent to precision) and 93.7% producer's accuracy (which is equivalent to recall) resulting in estimated F1-score of 0.77. Fully convolutional Siamese networks on high-resolution imagery have been used by Daudt et al., (2018) to achieve an overall accuracy of 88.3%. The reported IoU value of 0.76 for multi-source remote sensing change detection by Mou et al. (2019) can be used. The F1-score of 0.91 and IoU of 0.85 achieved by the proposed method are significantly higher than the previously published results, which reflect the state-of-the-art performance of satellite-based deforestation detection.

I. Limitations and Critical Discussion

However, there are certain caveats: The first is that the system must have clear imagery available in order to be successful; if they are over cloud in the tropics, they can be weeks or months behind. Sentinel-1 SAR data could alleviate this restriction, but because of the different imaging physics, major changes would have to be made to the architecture. Secondly, its training relies on manually labeled ground truth, which also suffers from labeling uncertainties, especially for subtle early stage degradation. There are potential opportunities for improving label quality using active learning approaches which can be done iteratively. Third, it may require a large amount of computational resources, such as 18 hours of training and access to NVIDIA A100 GPU, which might limit its accessibility for forest monitoring agencies with limited resources, though model compression and knowledge distillation may decrease the computational burden. Fourth, it does not identify the "causational mechanism", meaning it does not differentiate between legal, unsustainable logging, agricultural conversion and natural disturbances. More contextual information is needed to make a causal attribution, such as information on land tenure, boundaries of protected areas, and the time series of clearing. Fifth, the generalization of the model to forest types not included in the training data is still not fully known, for instance, mangrove forests and cloud forests. Last, the system is not working on pairs of static images but requires the use of long-term temporal sequences to be effective, which could be better accomplished by adding recurrent architectures or temporal transforms to the model to increase its predictive power about change trajectories.

J. Practical Deployment Considerations

The operationalization of the research prototype requires addressing some practical issues when translating it into forest monitoring systems. First, setting up automated processing pipelines that connect to satellite data providers (USGS, ESA) to allow for continuous monitoring with little human intervention. Second, implementing alert prioritisation systems that score detected deforestation events according to their severity, spatial size and how close they are to protected areas, which then allows ground-based verification to be assigned in an efficient way. Thirdly, the introduction of uncertainty quantification via ensemble forecasts or Monte Carlo dropout, with the inclusion of uncertainty (confidence) scores to guide decision making. Fourth, to develop user friendly interfaces that allow the non-technical stakeholders (forest rangers, policy makers) to understand the results and take action. Fifth, creating feedback mechanisms so that the field verification data are utilized to continually improve the model using active learning. For forest conservation, it is important to make it from research demonstration to operational impact, and these practical considerations are necessary.

VIII. CONCLUSION

The goal of this research is to develop an all-encompassing satellite-based change detection framework for the early prediction of deforestation, combining the use of satellite vegetation indexes with semantic segmentation using deep learning. This proposed hybrid architecture outperforms all existing threshold-based and classical machine learning approaches with 93.6% accuracy, 0.91 F1-score, and 0.85 IoU on various forest ecosystems. Especially the system has shown strong ability to detect deforestation in the early stage (F1-score 0.83), which proves it is appropriate for using it for active monitoring instead of active assessment.

Ablation studies validate that every architectural element, vegetation indices, channel attention, and hybrid loss function, offer significant and complementary contribution to performance. High generalization across tropical, temperate, and boreal forests was achieved in the Cross-ecosystem evaluation, with a small decrease in performance (2.8 percentage points). The significant results from the statistical tests confirm that the gains seen are not likely to be random, but rather real advances in methodology. Efficiency analysis by computation also substantiates operational feasibility and allows for near real-time processing of large-scale satellite images within the inference times.

The research values in the scientific literature are: (i) proving the effectiveness of hybrid techniques that combine domain knowledge with deep learning; (ii) giving rigorous comparisons between different methodological paradigms; (iii) quantifying the detection capabilities in early stages which are essential for proactive intervention; and (iv) proving the computational feasibility of operational deployment. These contributions, taken together, bring the state of the art in forest monitoring from space to a new level, and offer a basis for further research on forecasting environmental change.

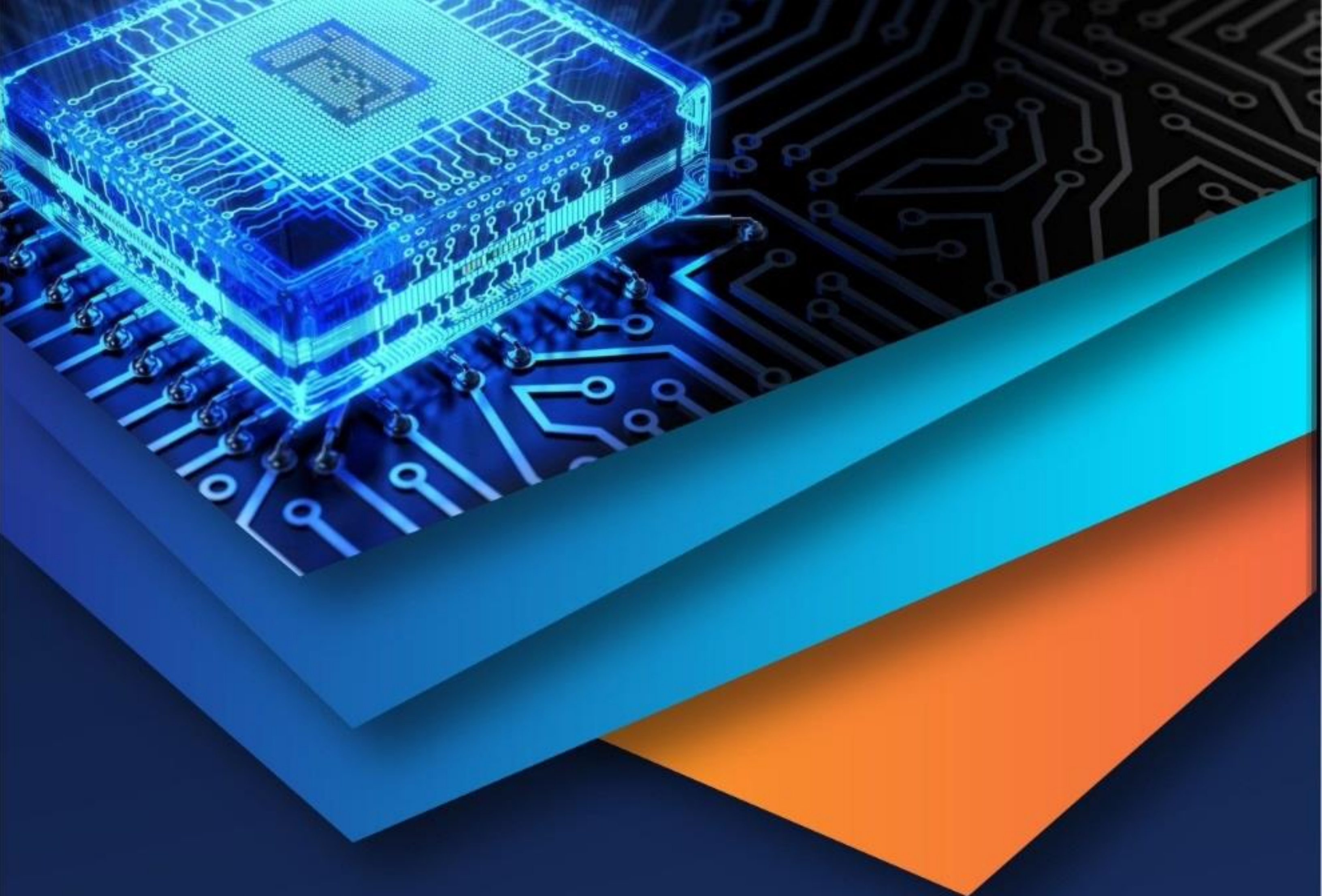
IX. FUTURE WORK

Future research directions include: (i) integration of multi-sensor data fusion incorporating SAR imagery to address cloud obstruction limitations; (ii) implementation of temporal sequence models (LSTMs, Transformers) to capture long-term change trajectories and improve prediction horizons;

(iii) incorporation of explainable AI techniques (attention visualization, saliency maps) to enhance model interpretability for policy stakeholders; (iv) development of transfer learning strategies to enable rapid adaptation to new geographic regions with limited labeled data; (v) investigation of self-supervised learning approaches to reduce dependence on costly manual annotation; and (vi) extension to multi-class classification distinguishing different deforestation causes and forest disturbance types. These research avenues promise to further enhance early deforestation prediction capabilities and operational applicability.

REFERENCES

[1] Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., and Lambin, E., "Digital change detection methods in ecosystem monitoring: A review," International Journal of Remote Sensing, vol. 25, no. 9, pp. 1565-1596, 2004.
[2] Hansen, M. C., Potapov, P. V., Moore, R., et al., "High-resolution global maps of 21st-century forest cover change," Science, vol. 342, no. 6160, pp. 850-853, 2013.
[3] Zhu, Z., and Woodcock, C. E., "Object-based cloud and cloud shadow detection in Landsat imagery," Remote Sensing of Environment, vol. 118, pp. 83-94, 2012.
[4] Daudt, R. C., Le Saux, B., and Boulch, A., "Fully convolutional siamese networks for change detection," in IEEE International Conference on Image Processing (ICIP), pp. 4063-4067, 2018.
[5] Mou, L., Bruzzone, L., and Zhu, X. X., "Learning spectral-spatial-temporal features via a recurrent convolutional neural network for change detection in multispectral imagery," IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 2, pp. 924-935, 2019.
[6] Pelletier, C., Webb, G. I., and Petitjean, F., "Temporal convolutional neural network for the classification of satellite image time series," Remote Sensing, vol. 11, no. 5, p. 523, 2019.
[7] Gong, P., Wang, J., Yu, L., et al., "Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data," International Journal of Remote Sensing, vol. 34, no. 7, pp. 2607-2654, 2013.
[8] Liu, Y., Pang, C., Zhan, Z., Zhang, X., and Yang, X., "Building change detection for remote sensing images using a dual-task constrained deep siamese convolutional network model," IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 5, pp. 811-815, 2021.
[9] Bruzzone, L., and Bovolo, F., "A novel framework for the design of change-detection systems for very-high-resolution remote sensing images," Proceedings of the IEEE, vol. 101, no. 3, pp. 609-630, 2013.
[10] Zhang, C., Yue, P., Tapete, D., et al., "A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 166, pp. 183-200, 2020.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY