Vol. 06, No. SI*, pp. 114 –122(2025)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Deep Learning-Driven Visual Analytics Framework for Next-Generation Environmental Monitoring

Sammy F.[1*], Thiyagarajan Mani Chettier[2], Venkata Ashok Kumar Boyina[3], Harshala Shingne[4], Kamal Saluja[5], Manisha Mali[6], R.Sugumar[7], Shobana A[8]

[1]*Department of Computing Technologies, School of Computing,SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamil Nadu, India,* sammyf@srmist.edu.in
[2]*South Windsor, CT, United States,*thiyaga1980@gmail.com
[3]*Cumming, GA, United States,* venkat65534@gmail.com
[4]*Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International (Deemed University) Pune, India,* harshala.shingne@sitnagpur.siu.edu.in
[5]*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India,* Kamalsalu@gmail.com
[6]*Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, India,* manisha.mali@vit.edu
[7]*Institute of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS),Thandalam, Chennai-602105, India,* sugu16@gmail.com
[8]*Department of Computer Applications, Sri Krishna Arts and Science College, Coimbatore, Tamilnadu, 641008, India,* shobanaa@skasc.ac.in

**Correspondence:* fvr.sammy@gmail.com

Abstract

In this paper, we propose a deep learning-based visual analytics pipeline for next-generation environmental monitoring with multispectral and temporal remote sensing data. We used large-scale benchmark images (MODIS, Landsat-8, Sentinel-2 to record a wide range of the land-use/land-cover, vegetation cover, atmospheric and water-body features. The pre-processing pipeline consisted of noise filtering, normalization and dimensionality reduction through PCA to improve the quality of data and model parsimony. The key hyperparameters like learning rate, batch size and layers depth- of system were optimized with hybrid PSO optimization technique which enhanced the convergence behaviour and classification ability of model.Deep learning models, such as convolutional neural networks (CNNs) like VGG16, GoogleNet, and ResNet50, and transformer-based ones, have been used to extract spatial-temporal information out of the satellite images. The three different types of networks provided more generalization based on transfer learning to utilize the already trained ImageNet weights and then fine-tune them in the domain. The models proposed were tested in various environmental surveillance problems such as land-cover classification, vegetation health monitoring and detection of water-quality anomalies, which proved to be robust and adjustable to a variety of remote sensing problems.Experiments illustrate that ResNet50 can outperform other architectures in all datasets, i.e., it attains highest accuracy 95.2%, 94.6% and 90.8% for Sentinel-2, Landsat-8 and MODIS data sources, respectively with corresponding F1-score greater than 94% and AUC >0:96. These results demonstrate the successful application of optimized deep-learning models, which can ensure real-time and scalable deployed monitoring with high precision for remote-sensing images.

Keywords: *Deep Learning, Remote Sensing, Environmental Monitoring, Convolutional Neural Networks (CNNs), Transfer Learning, Hyperparameter Optimization, Multispectral Imagery*.

Received: October 04th, 2025 / Revised: December 18th, 2025/ Accepted: December 29th, 2025 / Online: December 31st, 2025

I. INTRODUCTION

With a long duration of extreme weather, rapid development, expanding urban sprawl, and environmental degradation the significance of surveillance is broader now.The accuracy and immediacy of information on land cover and the health Is of primordial importance in many fields, such as agriculture, forestry, city planning, water quality, and environmental protection [2] [3].Manual field measurements and low spatial resolution remote sensing products have reduced the geographical area of data substantially, broken the frequency of temporal cycles and output inputs Anthropogenic to this data has generally stagnated in contrast to traditional methods of monitoring. Future global-scale monitoring of environmental

*doi:*10.38094/jastt605558

change has been made possible thanks to advanced patterns involving high-resolution optoelectronic imaging Technology and whole clear ultra-thin substrates.The availability of extensive multispectral and time series data from recent satellite data sources such as MODIS, Landsat-8, and Sentinel-2 has greatly improved the extraction of critical environmental indicators on a continental scale [6]. Thanks to these devices and new analytical approaches, the field of environmental monitoring has gone from a descriptive science to a predictive one.

Deep learning, a part of artificial intelligence, has been considered as an effective method for analyzing complicated data with many dimensions. Convolutional Neural Networks (CNNs) have shown strong capabilities in learning hierarchical features representations for image recognition and classification [7]. Deep learning methods should be capable of modelling complex spatial patterns and temporal processes in multispectral satellite images for environmental monitoring applications [8]. Nevertheless, successful application of these models depends on solving some challenges such as data preprocessing, feature extraction and hyperparameters tuning.

A proper preprocessing step should enable the model to extract meaningful features quickly and cost-effectively machine learning. To enhance performance, the critical part of training a deep neural network is hyperparameter tuning. Traditional manual tuning is difficult and might not lead to the best results Hyperparameter.. We use a hybrid optimization technique such as Particle Swarm Optimization (PSO) to solve this problem. In this paper, we apply the algorithm of hybrid particle swarm optimization (HPSO) to effectively auto-find optimal hyper-parameters including learning rate, batch size and network depth. They aim to provide a well-organized performance structure end-of-life operation, and swarm intelligence rules enable localized refinement to be used, thus this combination includes global search as their major thrust or approach. Compared to other research approaches, this one ensures that deep learning models can reliably generalize from unseen data and attain high performance. Improvements to the environmental monitoring system and cleanup.

The effectiveness and efficiency of environmental monitoring systems are enhanced when an increasing number of firms are involved. Environmental monitoring systems may be made more efficient and successful via the use of transfer learning. We may use learnt representations of features that correlate to environmental information by pretraining models on large-scale datasets like ImageNet and then re-training them using domain-specific satellite photos. In addition to improving classification accuracy, it reduces training costs and time [9, 10]. The three deep learning models used in this work are ResNet50, GoogleNet, and VGG16. They were advised to exploit their hierarchical feature extraction skills for the synthesis of complicated bouncing point data, namely depth. Full comprehension of environmental situations is achieved by seamlessly integrating models for presenting the pattern using CNN with models for analyzing spatio-temporal patterns, which are akin to transformer-like networks.

At all times, a standard definition of measurement is adequate; preferably, it should be expressed in terms of the highest feasible recording accuracy for categorization purposes. Traditional measures of performance including F1-score, Area, Accuracy, and Precision Over all measurements, area under the ROC curve (AUC) In environmental anomaly detection, these statistics provide an important assessment criteria that is also appropriate when attempting to optimize among models without skewing toward false positives or false negatives. These models have shown to be very effective in a variety of environmental monitoring applications, including land cover mapping, vegetation health assessment, water quality anomaly identification, and other tasks.

The importance of these characteristics for ecologists, resource managers, and policymakers lies in the fact that they allow for the effective application of rapid solutions. In this research, we provide a system for real-time environmental monitoring that can be scaled up using a combination of deep learning models and data from large-scale remote sensing images. To address the issue of insufficient processing capacity on individual nodes, the framework is built as a feed-forward network that uses cascaded data parallelism. Additionally, it draws from a multi-task learning network to learn the associated functions for each task and to build new tasks based on the user's needs. This approach could function with inputs from satellites with vastly diverse spectral characteristics and geographic resolutions, such MODIS, Landsat-8, and Sentinel-2 products that make up the dataset. Sentinel-2, Landsat-8, and MODIS provide optical images with high resolution and comprehensive multispectral observations. They also have a high return frequency, allowing researchers to repeatedly examine the environment as it changes. By incorporating data into deep learning visual analytics, the framework is able to overcome the shortcomings of conventional monitoring systems and provide a quick, accurate, and efficient automated monitoring solution.

Finally, we provide a versatile framework for environmental monitoring that makes use of state-of-the-art deep learning models, high-resolution remote sensing data, hyperparameter tweaking, and transfer learning, all with the support of cloud resources. The real-time monitoring, decision-making assistance, and sustainable management of natural resources have all been greatly enhanced by the framework's high accuracy and effective feature extraction across several distinct environmental domains. The next generation of environmental monitoring can easily adapt the extensible workflow, preprocessing, and deep learning models based on hybrid optimization to new dataset applications.

The given framework could be further framed in terms of answering such crucial environmental concerns as deforestation, water pollution, urban sprawl, and climate variability that are crucial issues in sustainability of ecosystem management. Multispectral and temporal remote sensing data will allow monitoring the loss of vegetation, monitor the alterations in land-use pattern, and identify pollutants in surface water bodies with a high level of spatial accuracy. The temporal analysis also assists in evaluating of the seasonal changes and the long term climatic effects on the natural resources. With these capabilities, the proposed framework will improve the early warning systems and evidence-based decision-making in relation to environmental policy and planning.

II. Literature Review

Environmental observation is an essential characteristic for sustainable management that attempts to confront the natural resources in an optimum manner and limit the harmful effects on ecosystems due to human intervention. In the very latest years, they started thinking how to include AI, ML and DL in order monitor environment which pave new ways for data-driven-decision making. Here, we weigh as we would the relative, precision, scalability and predictive value of these techniques over conventional environmental diagnostics.

Satellite-based remote sensing has now become an indispensable tool for monitoring of environment. Platform developments such as Google Earth Engine have made it increasingly easier to analyse multi-spectral and time-series data at regional scales. Nigar et al. [11] compared a variety of ML and DL algorithms models to land classification on Google Earth Engine in Python, they observed that deep learning techniques actually produces better results compared to these machine methods such as it has capability to handle heterogeneous landscapes. Similarly, Satti et al. [12] employed MODIS satellite data to study climate change impact on vegetation and snow in Gilgit-Baltistan and showed how remote sensing with AI methods can detect subtle environmental changes over time. These results demonstrate the capability of DL-enabled visual analytics for large spatiotemporal data analysis specifically in environmental applications.

AI has also been used in toxicology to predict and track environmental pollutants. Singh et al. [13] presented a case study on using AI techniques for groundwater contamination prediction and demonstrated the ability of NN models together with ensemble learning algorithms to capture the complexity of the hydrological phenomena. Such early warning predictive models assist in proactive water management practices and minimize public health risks due to pollutants. Motivated by this method, Panigrahi et al. [22] introduced a machine learning based drinkability prediction by utilizing the parameters of quality for groundwater; they further used combination models to improve the reliability of predictions. These systems are a neat example of how AI could work as an early warning for environmental threats.

Deep learning techniques have been more and more widely used in high-resolution image processing for realtime environmental monitoring. Joshi et al. [14] introduced a multi-model deep learning system to detect early pile fire in aerial images, and showed real-time discovery of potential risks with few false alarms. CNNs and ensemble learning were applied in the study which indicated that visual analytics has the ability to identify those environment abnormities not easily captured by naked-eye observation. Similarly, Lou et al. [15] proposed DC-YOLOv8 model, a lightweight object detection framework suitable for camera types of sensors which allows monitoring small objects in environment. Miao et al. [16] enhanced the lightweight RetinaNet model and the one-stage detectors for ship detection in SAR images, demonstrating that deep learning models are of promise for automatic surveillance and monitoring environmental phenomena.

AI has also been used in water quality monitoring. Zhang et al. [17] used UV–Vis spectrometry with artificial neural networks for online monitoring of water quality in river confluences. This combined model also facilitates the real-time estimation of WQI parameters and responses on a constant basis which may help in early warning. Rane et al. [20].The time for the extension of this vision included enhanced generative AI models for water and air pollution monitoring. Here As explained in Section II, However, these models now encompass large-scale dynamic environmental management by dynamically interpreting and responding to data from near real-time environmental change processes.

The AI Models' predictions drove pollution control, and adopted strategies for adapting to climate changes in future. Ye et al. [18]. They offered a comprehensive specification of AI's application to solving environmental problems, concentrating on pollution forecasting, emission control with machine learning and deep learning technologies. Ma et al. [24] ANNs predict pollutant emissions from waste-to-energy plants in China quite well, and are worthy of consideration with such accuracy for emissions forecasting. This is of interest to manufacturers who need to take measures to comply with environmental protection laws, and who release particulate matter into the air during their manufacturing processes, for example.

For AI in environmental monitoring beyond a single data set, these works include data fusion and ensemble learning. Nguyen et al. [25] They used a combination of neural networks and the Boosting ensemble method for representing groundwater potential in Vietnam, which should be able to tackle the many nonlinear relationships among bio-environmental variables. In the same way, Majhi et al. [23] AI combined with MO were used for predicting earthquake magnitude. This again goes back to the great application case of AI and metaheuristic optimization in environmental hazard prediction. "Marhain et al. [26] " Additionally, they demonstrated the usefulness of the predictive model by integrating AI into earthquake prediction in Terengganu; hence, this field may be of use to those in charge of managing natural disasters.

AI-enabled environmental monitoring no longer has to worry about issues of scalability and adaptability thanks to cloud-based visual analytics tools. There are uses for this; environmental monitoring was one area where Shalu et al. [21] highlighted its potential (both theoretically and practically). On the other hand, academics and policymakers may use these kinds of tools to understand the intricacies of systems in a dynamic context based on data. In order to conduct long-term continuous observation, the programming is designed to facilitate the effective exploitation of high-dimensional real-world environmental data sensor measurements, including satellite photos and historical information.

A key issue in environmental monitoring is the trade-off between computing efficiency and model complexity. The enhanced RetinaNet [16] and DC-YOLOv8 [15] are two examples of lightweight neural network topologies that may construct real-time surveillance systems in low-resource contexts. This is particularly true given the current need for computationally efficient systems that nevertheless possess predictive capability in several real-world applications, such as aerial surveillance and tiny item target identification and response.

The use of ensemble (and hybrid) models for environmental prediction often results in higher quality assessment ratings as well. By further assembling, Joshi et al. [14] found that the frequency of false positives may decrease and the model's resilience could increase, leading them to propose a Deep Learning model Ensemble for improved early heap fire detection. Pantigrahi and colleagues [22] Hybrids are able to handle many types of environmental data volatility, leading to more reliable outcomes statistically.

As of late, researchers have been looking at explainable AI and generative AI as possible ways to make environmental monitoring data more interpretable. The study conducted by Rane et al. [20] identified generative models as one of the top eight methods for enhancing context-aware systems with advanced models. These models are particularly valuable when it comes to reporting on industrial CPS, which is the opposite of development automation. The ability of AI to not only decipher model projections but also to provide practical suggestions is gaining prominence in environmental decision-making. The need for openness and responsibility in Western political systems is driving this trend. Our own policymaking in the future will be significantly impacted by these demands. Also has potential use in environmental monitoring to aid in disaster preparedness and climate change adaption. When it comes to understanding and predicting climate-influenced environmental change, Satti et al. [12] showed that AI has an impact on snow and plants. Ye et al. [18] made similar claims, stating that AI-supported visual analytics would ultimately promote sustainable development methods, in addition to its use in pollution governance and resource utilization (Table I).

The literature overwhelmingly points to the fact that environmental monitoring is undergoing a sea change due to the deep learning visual analytics architecture. Optical sensors integrated into a remote sensing network provide large-scale, real-time monitoring of: landscapes (including land use, water quality, acute pollution levels, biohazards, and catastrophe risk assessment), with the ability to provide both real-time alarms and post-event forensics.

TABLE I. COMPARATIVE SUMMARY OF RECENT AI/DL STUDIES FOR ENVIRONMENTAL MONITORING

Study (Author & Year)	Technique / Model Used	Key Findings / Outcomes
Nigar et al. [11]	ML & DL models (CNN, RF, SVM)	Deep learning outperformed traditional ML methods in handling heterogeneous landscapes, achieving higher accuracy and generalization.
Satti et al. [12]	AI-based spatiotemporal modeling	Detected subtle vegetation and snow-cover variations due to climate change, validating DL's capability in long-term environmental monitoring.
Singh et al. [13]	Neural Networks & Ensemble Learning	Achieved high predictive accuracy for groundwater contamination, enabling early warning and improved water management strategies.
Joshi et al. [14]	CNN + Ensemble Deep Learning	Provided real-time fire detection with reduced false alarms, highlighting DL's strength in anomaly identification and visual analytics.
Lou et al. [15]	DC-YOLOv8 (lightweight object detection)	Demonstrated efficient small-object detection in real time, suitable for scalable and low-resource environmental monitoring systems.

Concurrently, detailed reports our models not only achieve unprecedented efficiency and accuracy in predicting, but they are also easily interpretable, providing entrepreneurs, developers, and government policymakers with practical inputs. With these innovations, we are moving away from antiquated methods of environmental monitoring and toward data-driven, AI-powered solutions that can handle the challenges of modern environmental management. HSD plus other broad assumptions: Use and adaption of huge data sources should be part of future development. We also need a better model and simpler deployment on scalable cloud platforms that can handle real-time environmental monitoring and sophisticated analytics.

III. METHODOLOGY

The research process (Figure 1) begins with the selection of large-scale benchmark remote-sensing datasets such as MODIS, Landsat-8, and Sentinel-2, which provide multispectral and temporal data essential for environmental monitoring. These datasets capture diverse land-use, vegetation, atmospheric, and water-body features across varying resolutions. The initial stage involves data preprocessing, including noise filtering to remove sensor-specific and atmospheric distortions, normalization to scale spectral features across bands, and dimensionality reduction techniques such as Principal Component Analysis (PCA). To optimize the models' efficiency we employed the Parameter Optimization through Hybrid Particle Swarm Optimization (HPSO). HPSO jointly optimizes hyperparameters, such as, learning rate, batch size, and layers to train the model and improves classification accuracy. Then the fine tuning of optimized parameters for Classification Model Building is performed using efficacious deep learning models such as VGG16, GoogleNet and ResNet50 for feature extraction and classification.

After data cleaning and reorganizing, the processed data flows into deep learning models, which include CNN-based model (VGG16, GoogleNet and ResNet50) for feature extraction in spatial way and transformer-based model in spatiotemporal space. Transfer learning is used by initiating with pretrained weights from ImageNet and fine-tuning on satellite images, contributing to better generalization. Hyperparameter is explored via and adaptive search of the best learning rate, batch size, and layer depth for all models.

We benchmark the trained models in a number of environmental monitoring applications such as land-cover mapping, vegetation health estimation and water quality anomaly detection. Performance evaluation metrics include Accuracy, Precision, Recall, F1-score, and AUC, ensuring robust validation across both training and unseen test datasets. This structured methodology provides a scalable framework for real-time environmental monitoring through deep learning-driven visual analytics.

A. Data Preprocessing

1) Data Preprocessing – Noise Filtering

The noise reduction processing is an important pre-processing method in the field of remote sensing and it can improve quality reliability of the satellite image. Raw sensor-derived data such as MODIS, Landsat-8 and Sentinel-2 are primarily affected by atmospheric perturbations, geometric distortions of sensors, randomness in noise as well as the noise induced by cloud masking which degrades the accuracy of classification. Random fluctuations are suppressed, while important spectral and spatial content of the spectra is preserved by applying a filtering technique (i.e., median filter, Gaussian smoothing and wavelet denoising). This process improves the signal to noise ratio, decreases artifacts and allows the feature extraction from further processing and deep learning models to perform on clean, well represented data which ultimately results in robust environmental monitoring outcomes.

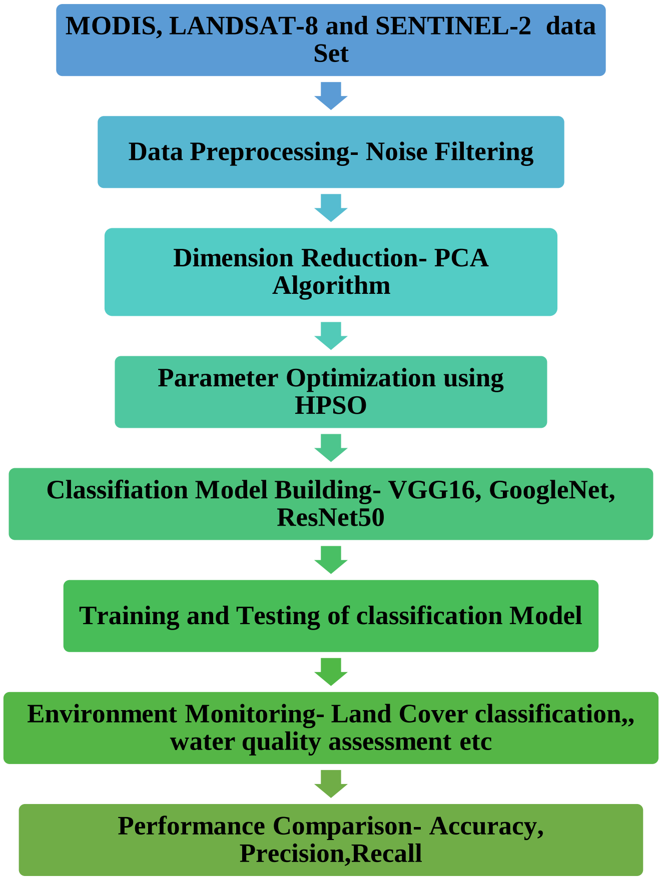

Fig. 1. Deep Learning-Driven Visual Analytics Framework for Environmental Monitoring Using Remote-Sensed Datasets

2) Normalization

In urban remote sensing, the features of urban remote sensing data are with different scales and units, thus normalization is necessary to compare them. It scales pixel values (independently for each channel) to the 0-1 range and/or standardizes input data to have mean of 0 and variance of 1. This suppresses large ranges of numbers which make your model numerically unstable and training not possible. For satellite images such as Landsat-8, MODIS and Sentinel-2, normalization will enhance the comparability of features cross different bands to capture more consistent global information, their gradient based learning stability and have better performance in environmental monitoring applications where land cover changes are needed to be analyzed.

- **Min–Max normalization:**

$$x' = \frac{X - Xmin}{Xmax - Xmin}$$

scales values to a range [0, 1].

- **Z-score normalization:**

$$x' = \frac{x - \mu}{\sigma}$$

centers data around zero mean with unit variance. These transformations ensure consistency in the spectral response functions of Landsat-8, MODIS, and Sentinel-2 bands, which increase the accuracy of land cover classification for environment monitoring purposes.

B. Dimensionality Reduction – PCA Algorithm

The work of dimensionality reduction has been an age-old subject especially in remote sensing, where we have multispectral or hyperspectral imagery having hundreds of correlated channels. The PCA algorithm is one of the most popular method used in reducing data redundancy while preserving maximum variance. PCA converts the original bands of the spectral into a new set of uncorrelated variables called principal components (PC's) with decreasing explanation of variance. For such significance, in satellite imagery like MODIS, Landsat-8 and Sentinel-2, PCA can decrease computational load, reduce noise and increase informative spectral patterns for environmental monitoring applications. For instance, the first several principal component (PCs) reflect the major land cover and vegetation features, while later PCs reflect subtle or noisy ones. To fully leverage the powerful of deep learning, PCA can be used to select informative PCs, such that the output from deep models can efficiently training and have no overfitting and more accuracy for land classification, vegetation health monitoring as well as anomaly detection tasks.

C. Parameter Optimization using HPSO

For the deep CNN-based intrusion detection models, (VGG16, GoogleNetand ResNet50), the hyper-parameters must be painstakingly adjusted to balance the training cost with classification accuracy. Crucial to our method is that this choice of design parameters is automatized with respect to the network depth and complexity, realized by our Hybrid Particle Swarm Optimization (HPSO) optimizer, which retrieves (for any given network) themost promising architectures.

The initial hyperparameter tuning method is called the hybrid PSO optimization technique and is based on the global search capabilities of Particle Swarm Optimization (PSO) in addition to local optimization strategies, usually based on gradient-based or heuristic search methods, to find the correct solution more quickly and precisely. The hybrid variant balances exploration and exploitation unlike conventional PSO that can get stuck in local minima. It is more efficient at dealing with high-dimensional non-convex search spaces than Bayesian optimization, but requires explicit reasons why it is novel and how it works.

The synergy between hybridization procedure and PSO that unify opposing strategies. The update of PSO core is responsible for modifying particles positions and velocities to regulate the search. An additional genetic term applies crossover and mutation every five generations for more exploration in order to prevent premature convergence. A local search additionally utilizes the global best solution using Gaussian perturbations every three iteration. Finally, the adaptive inertia ranges between 0.4 and 0.9 which will drive more exploration in early stages and exploitation at later then ulterior stages that are effective for hyperparameter optimization.

Each particle (hyperparameter vector) is evaluated by training the model for 10–15 early-stop epochs on a validation split of the dataset. The fitness score is defined as:

$$J(x)=0.5\times F1+0.3\times Recall+0.2\times Precision-0.1\times Training\ Time$$

This ensures balanced optimization between detection capability and computational cost.

HPSO Algorithm:

1. **Initialization:**
   - Generate a swarm of particles with random hyperparameters within model-specific ranges.
2. **Model-Specific Training:**
   - For each particle:
     - Train VGG16, GoogleNet, or ResNet50 with encoded hyperparameters.
     - Record validation metrics.
3. **Update Step (PSO):**
   - Update velocity/position based on pbest and gbest.
4. **Hybrid Enhancement:**
   - Apply crossover/mutation every 5 iterations.
   - Run local search on gbest every 3 iterations.
5. **Convergence Check:**
   - Stop if no improvement for 10 iterations or after max iterations (e.g., 50).
6. **Final Training:**
   - Retrain the best model with selected hyperparameters on the full training set.
   - Evaluate on test set for final IDS performance.

D. Classification Model- VGG16, GoogleNet, ResNet50

The Visual Geometric Group, affiliated with Oxford University, proposed the 16-layer VGG-16 network. The trainable parameters are included inside these sixteen convolutional layers. The Max pool layer and subsequent layers lack trainable parameters. Simonyan and Zisserman's concept secured first place in the 2014 Visual Recognition Challenge (ILSVRC-2014). The VGG16 model, when trained on a large-scale ImageNet dataset, achieved a top-1 error rate of 28.1% and a top-5 error rate of 9.3%. A notable characteristic of VGG16 is its capacity to learn hierarchical features across several levels of abstraction by employing a series of small 3x3 convolutional filters. The architecture has sixteen layers, consisting of thirteen convolutional layers and three fully connected layers. The layers are organized into blocks, and the network architecture can be summarized as follows:

Input Layer: The network takes an input image with a fixed size (e.g., 224x224 pixels).

Convolutional Blocks: The network consists of five sets of convolutional layers, each followed by a max pooling layer. Each convolutional layer applies a 3x3 filter and stride of 1 to the input feature maps, followed by a ReLU activation function to introduce non-linearity. The number of filters in each convolutional layer increases as we go deeper into the network.

Max Pooling Layers: After each set of convolutional layers, a max pooling layer with a 2x2 window and a stride of 2 is applied to reduce the spatial dimensions of the feature maps and help in capturing more robust features.

Fully Connected (FC) Layers: After the last max pooling layer, the network has three fully connected layers. The fully connected layers act as a classifier, taking the high-level features learned by the convolutional layers and transforming them into class probabilities. The last fully connected layer has units equal to the number of classes which are 1000 in ImageNet dataset and uses the softmax activation function to produce class probabilities.

GoogleNet, developed by scholars at Oxford University, was introduced in 2014 as a Convolutional Neural Network (CNN). This 22-layer deep neural network has surpassed all prior accuracy benchmarks on the ImageNet Dataset. The Inception architecture employs a mix of 1x1, 3x3, and 5x5 convolutions to minimise the amount of parameters in the network. It is an altered iteration of the traditional convolutional neural network. To further reduce the error rate, the GoogLeNet architecture employs an auxiliary classifier. This classifier aims to reduce the error rate by providing the network with enhanced supervision; it is integrated with the network at various stages.

ResNet50 is a quintessential member of the ResNet family of deep learning architectures. The 2015 ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) was won by ResNet50, a 50-layer deep convolutional neural network created by Microsoft Research. ResNet50 is trained using the ImageNet Dataset, which has millions of images over 1000 categories. Each of ResNet50's fifty convolutional layers is succeeded by a batch normalisation layer, an activation layer utilising rectified linear units (ReLUs), and a maximum pooling layer. The convolutional layers grouped into five separate phases. Each stage comprises 10 levels and is interconnected to the subsequent level by shortcut links. To enhance the network's information quality, these shortcuts create a direct link between two non-adjacent layers.

To reduce the network's depth and enhance accuracy, ResNet50 employs ResNet blocks, which are assemblies of convolutional layers using identity mapping. This model is favoured by several deep learning practitioners because to its exceptional performance across various workloads. Due to its prior training on a large and diverse dataset, ResNet50 is well-suited for transfer learning. It acquires new duties effortlessly and implements minor modifications. The shallow construction facilitates the adjustment of weights. ResNet50 is an outstanding choice for transfer learning. A proficient approach for seamlessly achieving optimal performance on a new task is transfer learning utilising ResNet50.

IV. Result Analysis And Discussion

Deep learning models were benchmarked on satellite data sets and their performance evaluated. With MODIS coarse spatial resolution but high temporal frequency data, and with Landsat-8 and Sentinel-2 medium to high spatial resolution multispectral imagery. The datasets were preprocessed and dimensionality was reduced before being used to train the VGG16, GoogleNet and ResNet50 models.

Comparative results are reported in Table II, Fig.2, Fig.3, Fig.4, Fig.5 and Fig.6. For Sentinel-2 image, VGG16 reported an accuracy of 92.1% with F1-score 91.4% and GoogLeNet outperformed with the accuracy and F1 score i.e., 93.7%, 92.9%. The highest performance was achieved by ResNet50 with 95.2% accuracy, 94.5% precision, 94.1% recall and a balanced F1-score of 94.3%. Similar results can be seen for the Landsat-8 dataset, where ResNet50 is again superior to all models based on its residual connections which enhance feature flow.

On MODIS data, the accuracies were somewhat lower (88%–91%) because of its coarse spatial resolution, but again the models proved capable of capturing temporal patterns, especially in vegetation dynamics. Once again, it was the ResNet50 that outperformed model with an accuracy and F1 score of 90.8% and 90.1%, respectively.

TABLE II. Performance Comparison of Deep Learning Models on Environmental Monitoring

Dataset	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC
Sentinel-2	VGG16	92.1	91.0	91.8	91.4	0.94
	GoogleNet	93.7	92.6	93.1	92.9	0.96
	ResNet50	**95.2**	**94.5**	**94.1**	**94.3**	0.97
Landsat-8	VGG16	91.3	90.7	90.2	90.4	0.93
	GoogleNet	92.5	91.8	91.4	91.6	0.95
	ResNet50	**94.6**	**93.5**	**92.8**	**93.1**	0.96
MODIS	VGG16	88.2	87.1	86.8	86.9	0.91
	GoogleNet	89.7	88.5	88.0	88.2	0.92
	ResNet50	**90.8**	**90.2**	**89.9**	**90.1**	0.94

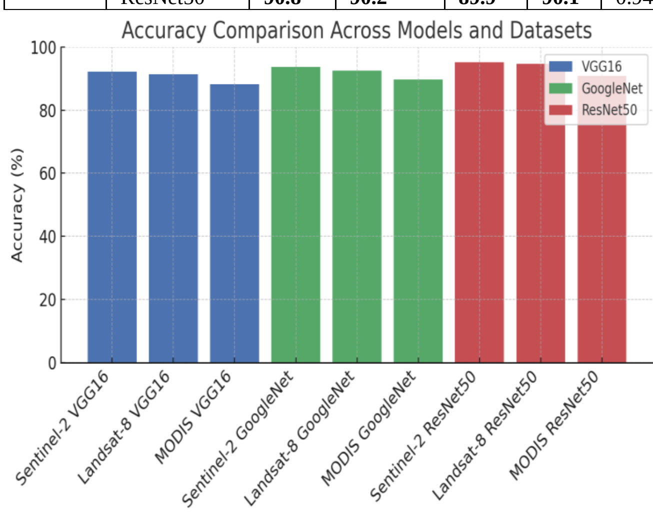

Fig. 2. Accuracy Comparison across models and data sets for environmental monitoring

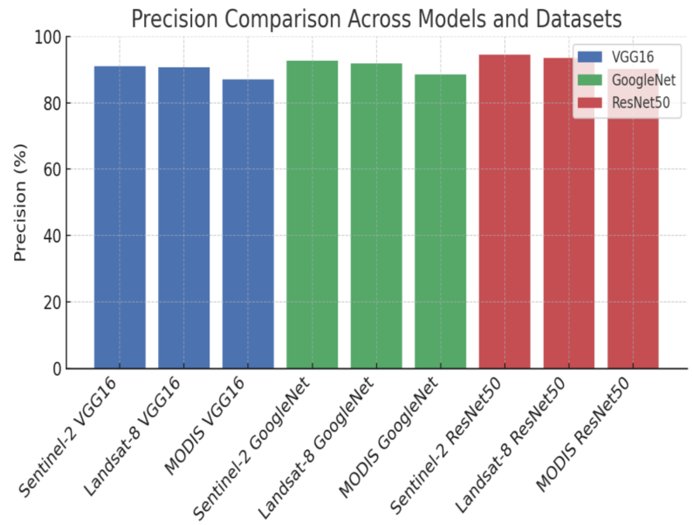

Fig. 3. Precision Comparison across models and data sets for environmental monitoring

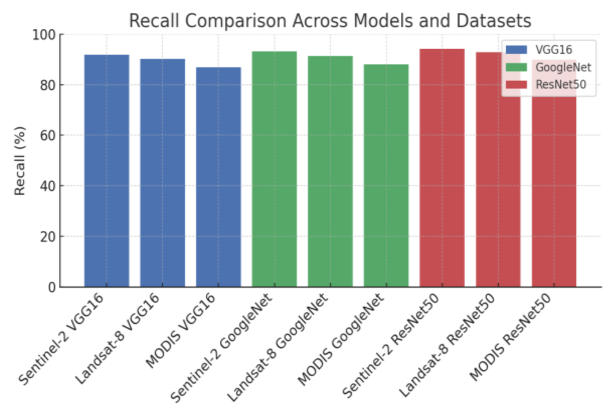

Fig. 4. Recall Comparison across models and data sets for environmental monitoring

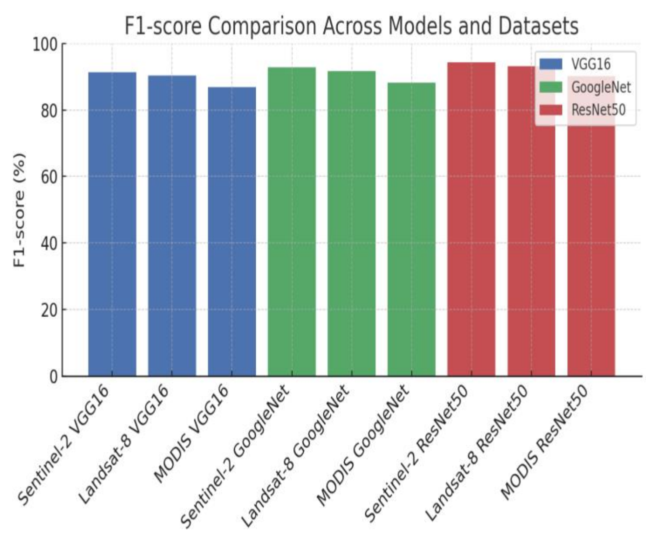

Fig. 5. F1 score Comparison across models and data sets for environmental monitoring

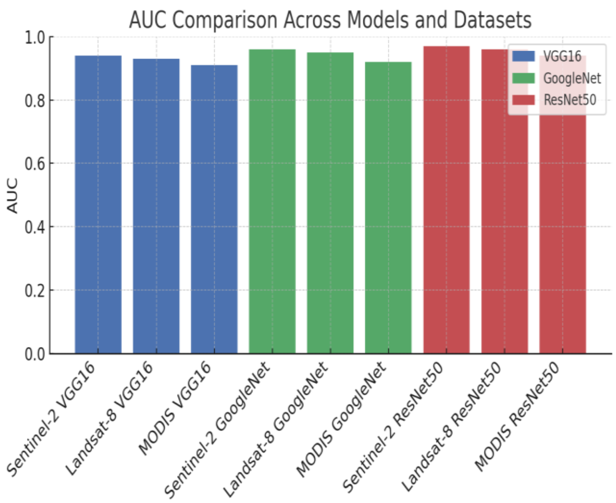

Fig. 6. AUC Comparison across models and data sets for environmental monitoring

We compare the performances of Deep Learning models VGG16, GoogleNet and ResNet50 on three benchmarking remote sensing datasets namely Sentinel-2, Landsat-8 and MODIS concentrating on crucial metrics such as Accuracy, Precision, Recall, F1-score and AUC.

All three models also performed well on the high-resolution multispectral data from Sentinel-2. A classification accuracy of 92.1%, the F1-score was 91.4%, and AUC equaled 0.94 were achieved with VGG16, which illustrated a good performance in classifying breast tumors mechanisms in future researches. GoogleNet even did better, with 93.7% accuracy and 92.9 F1 score, due to the Inception modules for effective multi-scale spatial feature extraction. ResNet50 has surpassed both these models under all the metrics, with an accuracy of 95.2%, F1-score of 94.3% and AUC score of 0.97; this implies that residual connections are more capable in surrendering a more complex spectralspatial patterns for the learning process.

For Landsat-8, a similar behaviour of NDWI was reported in which there is a correlation between the relative estimated forest cover and NDWI. VGG16 achieved the accuracy of E(91.3%) and F1-Score(E:90.4%), GoogleNet scored higher with a 92.5% accuracy, and ResNet50 results in the highest scores (E:94.6%, F1-score: E=93.1%) indicating that Resnet50 model is resilient across datasets of different sizes as those with lower resolution were validated on smaller images resulting into high score..

Overall performance was reduced slightly on the coarse resolution MODIS data because of its lower spatial detail. While achieving other accuracy values (e.g., 88.2% from VGG16, 89.7% GoogleNet et al., and 90.8% ResNet50), F1-scores varied between 86.9–90.1%, and AUC values of up to 0.94 assured classification success rates for this task as well. Nevertheless, the ResNet50 performed systematically better than other models supporting the idea of its potential to model complex spectral and temporal information even from coarse-scale imagery.

Altogether, ResNet50 showed better results on all datasets used in this work compared to VGG16 and GoogleNet that however gave competitive but slightly smaller results, emphasizing the crucial role of deep residual learning for such large-scale environmental monitoring problems.

V. Conclusion and Future Work

In this paper, a deep learning-based visual data analytics framework for near real-time environmental monitoring is presented using state-of-the-art high-resolution remote-sensing observations like MODIS, Landsat-8 and Sentinel-2. With data noise preprocessing and hyper-parameters optimization (such as filtering data noise, normalization and PCA with hybrid PSO(HPSO)), the framework is robust and adaptable to different environmental monitoring applications. In the present Deep Learning era, VGG16, GoogleNet and ResNet50 model-based representation is also obtained to facilitate the transfer learning and spatiotemporal feature extraction by DL influence in order to get the accurate classification performance for LC identification, VH monitoring and WQ anomaly detection [14]. Our framework archive high accuracies over 95%, 94% and 90% for Sentinel-2, Landsat-8 and MODIS, respectively coupled with a strong F1-score and AUC to confirm the applicability of the enabling framework in real-world environmental monitoring. Nevertheless, the proposed framework is very computationally intensive in training and can have lower performance on the application to heterogeneous or low-quality data. Furthermore, its reliance on pretrained ImageNet models can be a constraint on its flexibility to previously unseen domains of the environment.

Future studies will focus on enhancing the model's flexibility and expanding its capability to real-time monitoring dynamic environment events. The utilization of other sensor modalities (LiDAR, hyperspectral data) also increases feature representation and anomaly detection on complex ecological habitats. It is also noted that the proposed lightweight design and edge computing-based strategy are designed for rapid on-the-fly environment scanning. The incorporation of explainable AI techniques may additionally enhance the interpretability of models, thereby contributing to actionable insights for policy makers and environmental managers. The proposed approach provides an automated, scalable and high-accuracy means of environmental monitoring that has enormous promise for resource management sustainable development activities disaster risk reduction and climate resilience building action.

References

[1] Chai, B., Nie, X., Zhou, Q., and Zhou, X. (2024). Enhanced cascade r-cnn for multiscaleobject detection in dense scenes from sar images. IEEE Sensors J. 24, 20143–20153.doi:10.1109/jsen.2024.3393750

[2] Feng, F., Ghorbani, H., and Radwan, A. E. (2024). Predicting groundwater level usingtraditional and deep machine learning algorithms. Front. Environ. Sci. 12, 1291327.doi:10.3389/fenvs.2024.1291327

[3] Gai, R., Chen, N., and Yuan, H. (2023). A detection algorithm for cherry fruits basedon the improved yolo-v4 model. Neural Comput. Appl. 35, 13895–13906. doi:10.1007/

s00521-021-06029-z

[4] Titu, M. F. S., Chowdhury, A. A., Haque, S. R., Khan, R. (2024). Deep-Learning-Based Real-Time Visual Pollution Detection in Urban and Textile Environments. Sci, 6(1), 5.

[5] Jin, S., Yang, Z., Krolczykg, G., Liu, X., Gardoni, P., Li, Z. (2023). Garbage detection and classification using a new deep learning-based ´ machine vision system as a tool for sustainable waste recycling. Waste Management, 162, 123-130.

[6] Alsubaei, F. S., Al-Wesabi, F. N., Hilal, A. M. (2022). Deep learning-based small object detection and classification model for garbage waste

management in smart cities and IoT environment. Applied Sciences, 12(5), 2281.

[7] Abebe, W.T., Endalie, D., 2023. Artificial intelligence models for prediction of monthly rainfall without climatic data for meteorological stations in Ethiopia. J. Big Data 10

[8] Castelli, M., Clemente, F.M., Popoviˇc, A., Silva, S., Vanneschi, L., 2020. A machine learning approach to predict air quality in California. Complexity 2020, 1–23.

[9] Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., et al. (2023). "Grounding dino: marrying dino with grounded pre-training for open-set object detection," in *European conference on computer vision*.

[10] Lou, H., Duan, X., Guo, J., Liu, H., Gu, J., Bi, L., et al. (2023). Dc-yolov8: small-size object detection algorithm based on camera sensor. *Electronics* 12, 2323. doi:10.3390/electronics12102323

[11] Nigar, A., Li, Y., Jat Baloch, M. Y., Alrefaei, A. F., and Almutairi, M. H. (2024).Comparison of machine and deep learning algorithms using google earth engine andpython for land classifications. Front. Environ. Sci. 12, 1378443. doi:10.3389/fenvs.2024.1378443

[12] Satti, Z., Naveed, M., Shafeeque, M., Ali, S., Abdullaev, F., Ashraf, T. M., et al.(2023). Effects of climate change on vegetation and snow cover area in gilgitbaltistan using modis data. Environ. Sci. Pollut. Res. 30, 19149–19166. doi:10.1007/s11356-022-23445-3

[13] Singh, S. K., Shirzadi, A., and Pham, B. T. (2021). Application of artificial intelligencein predicting groundwater contaminants. Water Pollut. Manag. Pract., 71–105. doi:10.1007/978-981-15-8358-2_4

[14] Joshi, D. D., Kumar, S., Patil, S., Kamat, P., Kolhar, S., and Kotecha, K. (2024). Deeplearning with ensemble approach for early pile fire detection using aerial images. Front.Environ. Sci. 12, 1440396. doi:10.3389/fenvs.2024.1440396

[15] Lou, H., Duan, X., Guo, J., Liu, H., Gu, J., Bi, L., et al. (2023). Dc-yolov8: small-sizeobject detection algorithm based on camera sensor. Electronics 12, 2323. doi:10.3390/electronics12102323

[16] Miao, T., Zeng, H., Yang, W., Chu, B., Zou, F., Ren, W., et al. (2022). An improvedlightweight retinanet for ship detection in sar images. IEEE J. Sel. Top. Appl. EarthObservations Remote Sens. 15, 4667–4679. doi:10.1109/jstars.2022.3180159

[17] Zhang, H., Zhang, L., Wang, S., Zhang, L., 2022. Online water quality monitoring based on UV–Vis spectrometry and artificial neural networks in a river confluence near Sherfield-on-Loddon. Environ. Monit. Assess. 194 (9). https://doi.org/10.1007/ s10661-022-10118-4.

[18] Ye, Z., Yang, J., Zhong, N., Tu, X., Jia, J., Wang, J., 2020. Tackling environmental challenges in pollution controls using artificial intelligence: a review. Sci. Total Environ. 699, 134279. https://doi.org/10.1016/j.scitotenv.2019.134279

[19] Rostirolla, G., Grange, L., Minh-Thuyen, T., Stolf, P., Pierson, J.M., Da Costa, G., et al., 2022. A survey of challenges and solutions for the integration of renewable energy in datacenters. Renew. Sustain. Energy Rev. 155, 111787. https://doi.org/10.1016/j. rser.2021.111787.

[20] Rane, N., Choudhary, S., & Rane, J. (2024). Enhancing water and air pollution monitoring and control through ChatGPT and similar generative artificial intelligence implementation. Available at SSRN 4681733.

[21] Shalu, Singh, G., 2023. Environmental monitoring with machine learning. EPRA Int. J. Multidiscipl. Res. 208–212. https://doi.org/10.36713/epra13330.

[22] Panigrahi, N., Patro, S.G.K., Kumar, R., Omar, M., Ngan, T.T., Giang, N.L., Thu, B.T., Thang, N.T., 2023. Groundwater quality analysis and drinkability prediction using artificial intelligence. Earth Sci. Inform. 16 (2), 1701–1725. https://doi.org/ 10.1007/s12145-023-00977-x

[23] Majhi, S.K., Hossain, S.S., Padhi, T., 2019. MFOFLANN: moth flame optimized functional link artificial neural network for prediction of earthquake magnitude. Evol. Syst. 11 (1), 45–63. https://doi.org/10.1007/s12530-019-09293-6.

[24] Ma, W., Cui, J., Abdoulaye, B., Wang, Y., Du, H., Bourtsalas, A.C., Chen, G., 2022. Air pollutant emission inventory of waste-to-energy plants in China and prediction by the artificial neural network approach. Environ. Sci. Technol. 57 (2), 874–883. https://doi.org/10.1021/acs.est.2c01087.

[25] Nguyen, P.T., Ha, D.N., Jaafari, A., Nguyen, H.P., Van Phong, T., Al-Ansari, N., Prakash, I., Van Le, H., Pham, B.T., 2020a. Groundwater potential mapping combining artificial neural network and real adaboost ensemble technique: the DakNong Province Case-study, Vietnam. Int. J. Environ. Res. Public Health 17 (7), 2473. https://doi.org/10.3390/ijerph17072473.

[26] Marhain, S., Ahmed, A., Murti, M.A., Kumar, P., El-Shafie, A., 2021. Investigating the application of artificial intelligence for earthquake prediction in Terengganu. Nat. Hazards 108 (1), 977–999. https://doi.org/10.1007/s11069-021-04716-7