Weight Optimization for Missing Data Prediction of Landslide Susceptibility Mapping in Remote Sensing Analysis

[1]Kanchana S, [2]Jayakarthik R, [3]Dineshbabu V, [4]Saranya M, [5]Srikanth Mylapalli and [6]Rajesh Kumar T
[1]Department of Computer Science, Faculty of Science and Humanities, SRM Institute of Science & Technology, Kattankulathur, Tamil Nadu, India.
[2]Department of Computer Science and Intelligence, Saveetha College of Liberal Arts and Sciences, Chennai, Tamil Nadu, India.
[3]Department of Information Technology, Karpagam Institute of Technology, Coimbatore, Tamil Nadu, India.
[4]Department of Computer Science and Engineering, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.
[5]Koneru lakshmaiah Education Foundation, vaddeswaram, Guntur, Andhra Pradesh, India.
[6]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Technical and Medical Sciences, Saveetha University, Chennai, Tamil Nadu, India.
[1]kanchans@srmist.edu.in, [2]drrjayakarthik@gmail.com, [3]dineshbabukit@gmail.com, [4]sm2317@srmist.edu.in, [5]mylapallisk@gmail.com,[6]t.rajesh61074@gmail.com

Correspondence should be addressed to Jayakarthik R : drrjayakarthik@gmail.com.

Article Info
Journal of Machine and Computing (http://anapub.co.ke/journals/jmc/jmc.html)
Doi: https://doi.org/10.53759/7669/jmc202404043
Received 15 May 2023; Revised from 17 October 2023; Accepted 08 January 2024.
Available online 05 April 2024.

Abstract – To keep track of changes to the Earth's surface, extensive time series of data from remote sensing using image processing is required. This research is motivated by the effectiveness of computational modelling techniques; however, the problem of missing data is multifaceted. When data at numerous a-periodic timestamps are absent during multi-temporal analysis, the issue becomes increasingly problematic. To make remote sensing time series analysis easier, weight optimised machine learning is used in this study to rebuild lost data. Keeping the causality restriction in mind, this method makes use of data from previous and subsequent timestamps. The architecture is based on an ensemble of numerous forecasting modules, built on the observed data in the time-series order. Dummy data is used to connect the forecasting modules, which were previously linked by the earlier half of the sequence. After that, iterative improvements are made to the dummy data to make it better fit the next segment of the sequence. On the basis of Landsat-7 TM-5 satellite imagery, the work has been proven to be accurate in forecasting missing images in normalised difference vegetation index time series. In a performance evaluation, the proposed forecasting model was shown to be effective.

Keywords – Missing Data, Weight Optimization, Machine Learning, Spatial temporal analysis, Landslide Susceptibility Mapping, Image Processing.

I. INTRODUCTION

Researchers have been able to access high-dimensional, multiresolution information thanks to enhanced remote sensing technologies. Satellite imaging can provide information on the Earth's surface at a variety of spectral, geographical, and temporal resolutions. Unmanned aerial vehicles with cameras installed on the rotor blades cannot produce such photos. In order to extract geographic information from these data sets, they are utilised to create maps that explain the topography, land cover, lithology, etc. Earth observation and administration, for example, are made easier as a result [1]. Satellite imaging is highly specialised and offers more information than simple photos. With the use of these cameras, you may see more than you could by simply standing still because they take pictures of vast areas. Maps of coastal ecosystems and assessments of land cover and ecosystem change have both benefited substantially from the use of satellite images. Remote sensing technologies are becoming more important and sought after for use in managing coastal ecosystems like estuaries, wetlands, and coral reefs as a result of improvements in sensor design and data analysis methodologies. Pictures have greatly aided in the mapping of coastal ecosystems as well as the assessment of land coverage and ecosystem change [2,

3]. Because of advancements in sensor design and data analysis approaches, remote sensing technologies are becoming increasingly relevant and sought for use in managing coastal ecosystems such as estuaries, wetlands, and coral reefs.

The amount of biotic and abiotic floating particles, as well as the composition of the dissolved compounds in the coastal waters, are all monitored by multispectral and hyperspectral sensors. Remote sensing technologies are becoming more important and in demand for use in managing coastal ecosystems like estuaries, wetlands, and coral reefs as a result of improvements in sensor design and data analysis methodologies. Geographically complicated coastal ecosystems exist. A successful field data gathering sampling method including ships, navigating markers, and survey equipment is necessary to accurately measure and validate remotely sensed data. Multispectral and hyperspectral sensors are used to monitor coastal land cover, dissolved compounds in coastal waters, and amounts of biotic and abiotic floating particles [4, 5]. Coastal environments are geographically complex. To verify and precisely measure remotely sensed data, an effective field data gathering sampling strategy utilising ships, navigating markers, and survey equipment is required [6, 7]. Earth surface dynamics can be studied using remote sensing data. The resulting data sets cover a wide range of space and time. It might be challenging to foresee the occurrence of landslides. Failure to do so, however, can have disastrous consequences, resulting in unwelcome catastrophes including property loss, community dislocation, and human casualties. By locating landslide-prone locations, research aims to prevent such calamities. The effectiveness of computational modelling techniques in similar crisis scenarios inspires this effort to investigate such modelling for LSM.

The fact that they frequently have low quality or missing values makes them less desirable. Clouds appear to be obscuring the Earth's surface in some optical satellite sensor data sets. Therefore, its usefulness is restricted. In some cases, for example, the ability to accurately track changes in vegetation over time can be hindered (see Arctic [8], Amazon [9], and general [10]). It is common for low-quality and missing variables to be eliminated range of "gap filling methods" (also known as "imputation approaches") [11, 12]. Only a few predictive algorithms have attempted to take advantage of the spatio-temporal character of the data [13]. A map of landslide susceptibility categorises areas into low, moderate, and high susceptibility areas. Both the positions and reasons of the collapses are taken into account on the landslide hazard map (slope, soil type and the impact of the flow of water in an area). A landslide's susceptibility is determined by the local terrain and the possibility that one will occur there. It forecasts "where" landslides will probably happen. For current environmental science, multi-year historical remote sensing data and ongoing periodic measurements are essential instruments. Understanding and predicting Earth system behaviour, identifying risky and anomalous processes, supporting economic activity and sustainable development are all supported by the data. [14, 15]. As of now, the satellite remote sensing business is well established, offering thousands of datasets as well as thousands of software tools and algorithms that cover a wide range of issues. Due to cloud cover, system failures, and other circumstances, missing data (data gaps) may affect the quality and usability of data. In high latitudes, cloud cover can result in data loss of up to 80% [16]. MODIS team delivers multiday using multi-temporal averaging data, for example, to deal around such problems [17]. This method is useful for studying long-term processes, but it is ineffective when applied to operational situations or processes that have a strong daily dynamic. For ecological [18], hydrological [19], meteorological [20], geological [21] and other forms of environmental study, land surface temperature (LST) is an essential geophysical parameter. To get a better understanding of the life processes, habitats, interactions, and biodiversity of species is the primary goal of an ecologist. Understanding the distribution of biotic and abiotic components in the ecosystem is the main objective of ecological. which includes both living and non-living elements and their interactions with one another. The study of water's occurrence, distribution, movement, characteristics, and interactions with the environment during each stage of the hydrologic cycle is known as hydrology. The study of the atmosphere, atmospheric phenomena, and the impact of the atmosphere on our weather is known as meteorology. The science of the atmosphere is concerned with the history of the earth and its life, particularly as preserved in rocks. It is the gaseous layer of the physical environment that surrounds a planet.

Over the last few years, there has been an explosion of research into machine learning algorithms and the methodologies that can be used to implement them. Models can be quickly created because many ML algorithms are freely available (mainly in Python and R [22]. Because of the machine The model utilised by machine learning has undergone a significant transformation from the previous model in recent years due to knowledge reformation and the rapid growth of neural network technology [23, 24]. Computers are used to process images, and there are two types of image processing: quantitative word image processing and analogue image analysis technology. It can also be used for picture enhancing, conversion, recovery, analysis, and other things, depending on how different approaches are used. However, there are numerous other data formats used in photo processing. This study shows the usage of available data from future timestamps, despite the fact that they violate the causality constraint. The proposed model includes a number of modules that use sequential data (images) from earlier time moments to predict the subsequent missing data (pictures) in the sequence. Each projected data item is adjusted iteratively to better fit the subsequent element in the entire sequence [25, 26]. Depending on how various methodologies are employed, it may also be used for image enhancement, conversion, recovery, analysis, and other purposes. However, photo processing involves a wide range of data formats [27, 28]. The use of available data from future timestamps, even if they contravene the causality requirement, is demonstrated in this paper. There are several modules in the proposed model that employ sequentially (pictures) from previous time moments in order to forecast the next missing data (pictures) in the sequence. In an iterative process, each of the predicted data items is fine-tuned to best fit the next component of the full sequence. We also use forecasting to keep the causality constraint during the data tuning procedure. When tuning, all projected data at intermediate time instants is available data in the sequence to forecast data

for the most recent available timestamp. Data for the final timestamp is compared to the original data for that timestamp to see if the predictions hold. The procedure continues until the accuracy of the predictions achieves a certain degree of contentment. The information from the future that is accessible is solely used to verify that the predicted image is consistent with the actual one. Section 2 lists the relevant works. Section 3 provides a breakdown of the model as presented here. Experiments to test the proposed model's efficiency are detailed in Section 4. Section 5 provides the final words on the matter.

II. LITERATURE REVIEW

The Chen, et al. [29] developed a geographically and temporally weighted regression model based on geo-statistical approaches to rebuild a target scene's missing data. The NDVI index uses reflected light in the visible and near-infrared ranges to recognize and quantify the existence of live, vegetated areas. Simply put, NDVI is a measure of how green and healthy the vegetation is in each pixel of a satellite image. We must grasp the electromagnetic spectrum in order to comprehend plant health. The ability to judge a plant's health based on how it reflects light and energy is essential to how NDVI works.A moving offset method was presented by Padhee and Dutta [30] for restoring missing values in time-series NDVI by utilising linear regression coefficients with a reference NDVI at another pixel with same conditions. Reconstruction of cloud-covered pixels can also be accomplished using multi-temporal approaches.

An efficient approach to handling missing values is the emphasis of Rani, [31]. A systematic approach to identifying missing values has been developed by the authors. The researchers used a dataset from the University of California, Irvine's (UCI) repository to conduct their research. There are three ways in which missing values can be inserted: random, MISSHASH, and MISSFIB. Data were examined with the use of) classifications for KNN, multivariate imputation (MICE), mean, and mode imputation. One of the two primary performance metrics for a regression model is the Root Mean Squared Error (RMSE). It calculates the standard deviation of values between those predicted by a model and those actually observed. It gives an estimate of how effectively the model can forecast the desired value (accuracy). Root Mean Square Error (RMSE) was used to compare classifiers in order to discover the optimal approach for imputation. The MICE imputation method fared better than other imputation methods, according to the study. In addition, its accuracy is not dependent on the classifier or the distribution of missing values.

Inductive Learning Algorithm (ILA4), developed by Elhassan, [32], is designed to deal with datasets containing missing values. ILA4 is based on a succession of ILA algorithms that have been improved to handle missing data. Datasets with varied degrees of completeness are subjected to ILA4 and compared to other well-established methods for dealing with datasets that contain missing values. Many existing ways to treating missing data, such as the Most Common Value (MCV) algorithm, the Most Common Value Restricted to a Concept (MCVRC) algorithm, and the Delete strategy, outperformed ILA4 in the majority of circumstances. Logistic regression, NB and random forest were tested against ILA4 to see if its accuracy was equivalent to or greater than theirs. ILA4 came out on top in all three tests.

A deep neutral network is a sort of technology that, like the human brain, comprises many levels and layers of neurons. The network as it is now consisting of an input layer, an activation function, and one or more hidden nodes. While striving to learn first from data fed into it, the network generates predictions [33, 34]. According to Park [35] a better ensemble model for disease prediction based on laboratory test results was created by combining two DNN (deep neural network) models with two machine learning (ML) models. The selection of 86 laboratory tests from datasets was based on criteria such as value counts, clinical significance, and missing values. This study collected 326,686 laboratory test results from 5145 patients. For 39 disorders, we used ICD-10 codes from the International Classification of Disorders, 10th revision (ICD-10). As a result of the data in these datasets, the TensorFlow-based DNN and the LightGBM and XGBoost machine learning models were built. For the five most frequent diseases, the optimised ensemble model had an F1 score of 81% and a prediction accuracy of 92%. The prediction power and disease categorization patterns of the deep learning and ML models differed. In order to determine which features were most important, we employed a confusion matrix and the SHAP value approach. Through the classification of diseases, our new ML model was able to achieve great efficiency in disease prediction. Diseases can now be predicted and diagnosed more accurately thanks to this research.

SFE+ is proposed in Zhu [36] as a sparse feature extraction network for LSP. As a starting point for the model, frequency ratios of environmental factors are determined for each of the landslides and environmental factors that affect them. Spatial data is one of the methods typically utilised to obtain physical data for GIS integrating. Remote sensors acquire information from earthly objects despite conducting direct touch. A geographic information system is a computer programme that analyses and displays geographically connected data (GIS). It uses knowledge related to a certain location.As a second step, the input data are sent via the input dropout. Then, the features are carried on through k percent lifetime sparsity layer, which is a hidden layer of the algorithm. Using hidden layers, these factors can be further reduced in order to get as many independent and redundant prediction features as feasible. Classifiers are used to implement the LSP in the research area. There are SVM, LR, and SGD models constructed using the SFE-support vector machine (SVM) algorithm.

Today, the sheer volume of remotely sensed data places it in the big data category. In order to interpret and make sense of distant sensing data, Sedona et al. [37] claim that (conventional) techniques. This is because a correctly designed ML work might entail a lot less data than the latter, which requires enormous quantities of training instances including millions of model parameters. Shirzadi et al. research [22] has also demonstrated that less complex ML can still be used in a variety of applications despite the deep learning trend. Data can be used to train supervised ML algorithms to build models for the

prediction of particular phenomena [38] as described by [39]. In another study [40], In addition to random forest and CART, boosted regression tree (BRT) and GLM were employed to identify areas at risk for landslides (classification and regression tree). According to their research, GLM was just as economical as RF and GBM (BRT) in terms of performance for LSM enough with other cutting-edge machine learning techniques. Virtual speakers are devices that use technique to develop animated facial images that can communicate [41]. With advancements in wireless sensor networks, the adoption of IoT devices is increasing daily. All IoT devices are connected to create a heterogeneous network that is extremely susceptible to outside threats [42].

III. PROPOSED METHODOLOGY

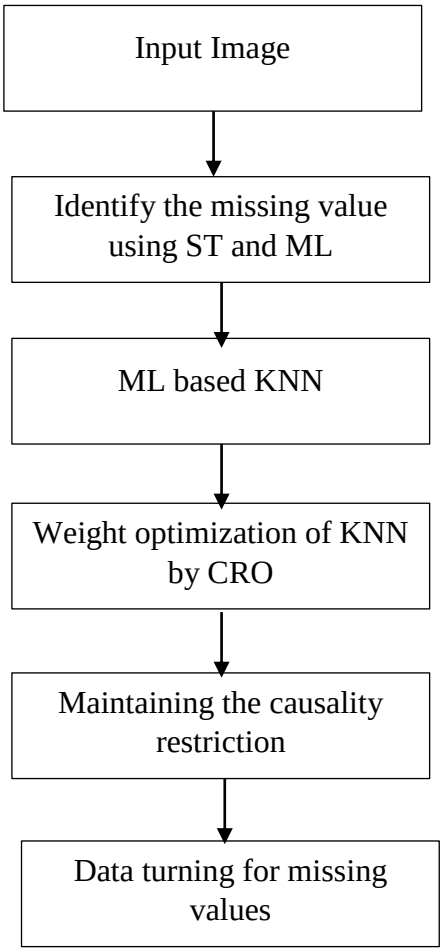

Fig 1. Working Flow of Proposed Model.

Problem Statement and Contributions

The following is a formal definition of the overarching issue addressed in this paper: In the case of a sequence of derived remote sensing imagery $I_1, \ldots, I_{m-1}, ?, I_{m+1}, \ldots, I_{x-1}, ?, I_{x+1}, \ldots, I_t$, where the symbol "?" denotes the missing images of certain a-periodic, the task is to forecast or rebuild the missing photos. The following are the contribution of the paper:

- Proposing weight optimization-based machine learning (ML) for remote sensing analysis using available Saptio-Temporal (ST) data from both earlier and later timestamps to predict missing data.
- Investigating ST analysis methods based on machine learning for missing picture reconstruction.
- This is where the causality restrictions in spatial time-series analysis are maintained.
- Based on Landsat-7 raw satellite time series [43], we tested the suggested technique in zones in India, each including several thousand pixels of missing NDVI images.

Weight/hyper-parameters optimization (HPO) in Machine Learning

At the outset, we must identify the important hyper-parameters that people must tweak to fit ML models into specific situations or datasets in order to improve them via HPO. There are supervised and unsupervised algorithms for machine learning models, depending on whether they are designed to model labelled or unlabeled datasets. An algorithm that uses labelled data to train on input features is referred to as supervised learning. In this study, supervised machine learning (ML) is employed to make predictions.

Supervised Learning Algorithms

In a predictive model function $f*$ is required to minimise the cost function $L(f(x_i), y_i)$, which models the discrepancy between the estimated output and ground-truth labels. As a result, there is no single method for creating a predictive model that works for everyone. Hyper-parameter settings restrict f's domain to a set of functions F in a limited number of model designs. As a result, this approach can be used to identify the best possible prediction model [44]:

$$f^* = \arg \min_{f \in F} \frac{1}{n} \sum_{i=1}^{n} L(f(x_i), y_i) \quad (1)$$

To calculate a cost function, x i is the feature vector for each sample, and L is its cost value for the nth training data point.

The other loss functions occur in supervised learning methods [45]. Hyper-parameter setups in ML algorithms generate a variety of predictive model topologies. This work, which makes use of the K-Nearest Neighbour (KNN) method, is one of several examples of supervised machine learning.

KNN

KNN is a simple machine learning technique that uses distances between data points to classify them. Using KNN, the predicted class of each test sample is allocated to the class to which the majority of its k-nearest neighbours belong in the training dataset.

T = x_1, y_1, y_2, x_n, y_n, where T is the training set of remote sensing pictures; the feature vectors of an instance and its class are denoted by ($x_i = c_1, c_2, \ldots, c_m$) for test instances; $I = (1; 2; \ldots, n)$. [46]:

$$y = \arg \max_{c_j} \sum_{x_i \in N_k(x)} I(y_i = c_j), i = 1,2, \ldots n; j = 1,2, \ldots, m, \quad (2)$$

As an indicator function, $N\ k\ (x)$ ($y_i = c_j$) is used to determine which of x's neighbours are the closest to it [47]. In KNN, the sum of nearest neighbours considered is the most important hyper-parameter. Over-fitting and excessive computational time will be required if k is too large or under-fitting is too small. Additional options include 'uniform' for the weighted function used in the prediction, depending on specific problems. Minor improvements can be made by adjusting parameters such as distance metrics and the Minkowski metric's power parameter. Coral Reef Optimization (CRO) is used to determine the ideal distance in this suggested metric. The K-NN algorithm places the new instance in the category that resembles the current subcategories the most, presuming that the new case and the previous cases are equivalent. After storing all the previous data, a new data point is categorised using the K-NN algorithm based on their similarity. This indicates that new data can be reliably and quickly categorised using the K-NN approach. The K-NN technique can be used for recovery even though classification algorithms are where it is most typically applied.

Coral Reefs Optimization Method

The corals that the algorithm generates at each stage are seeking to cling to the reef. The process is based on the growth of coral reefs and coral species. The coral reproduction is repeated several times with different operators to produce better outcomes. The three methods of sexual reproduction used by corals are asexual reproduction, broadcast spawning, and asexual reproduction (budding or fragmentation). It is a meta-heuristic programme that replicates the growth of corals in the reef. The corals of this method are made up of a variety of distinct coded solutions. In the reef, these corals are created at random and placed on a grid. Each of the grid's empty cells is filled at random with the corals of the starting population. There should be fewer solutions than the number of square grid cells, so that new corals can grow in the reef at a later date. It is also necessary to have a health function to check on the quality of the solutions and to come up with better ones. The objective function of optimising health can be the same as this subjective function. The method is based on the reproduction of corals and the creation of coral reefs. To get better results, the coral reproduction is repeated many times with a variety of operators. asexual reproduction, broadcast spawning, and asexual reproduction are the three types of sexual reproduction that corals use (budding or fragmentation). The corals that are generated in each stage of the algorithm are attempting to attach themselves to the reef. How well the solution optimises the problem, or how often it is to discover an empty spot, is a factor in this step's success. Internal and external reproduction operators develop solutions that are applied to the reef. For each new solution, the health function's value is determined.

The steps of CRO technique are as shadows.

Initialization:

As a first step, a reef is constructed from a matrices R with N rows and M columns. The generated population of solutions is then randomly distributed throughout the matrix's cells. Only one solution can be found in each cell. This is because if every cell in the matrix is already filled, new solutions will have a hard time finding a space and growing. For this, the ratio of R's vacant cells to its filled cells is calculated when each random solution is generated and placed in the matrix. The process of creating the initial population will be halted if this ratio falls below 0.4. For a 10 x 10 matrix with 100 cells, there are 72 random solutions (28/72 0:4). A unique identification is assigned to each solution after it has been generated

and added to a list. The R-matrix contains an identification that is randomly assigned to one of its cells. In addition, the procedure is used to determine the health function value of any solution generated.

External reproduction operator:
The solutions are applied to this operator in two stages. First, a subset of the matrix's solutions is chosen at random. Fb is a user-specified parameter that specifies the percentage of solutions that should be selected. When it comes to performing an external reproduction operator, reproduction can be done to distinct pairs of solutions if the number of selected solutions is even. You'll see your options for this phase separated into two lists: those you've picked and those you've left unselected. The cross-over operator is applied to each pair of separated solutions in the list of selected solutions in the second phase, using the roulette wheel approach. This step generates new solutions. If the cross-over operator is used, each pair of solutions has three distinct random points (ranging from 1 to n-1). The four pieces of the answer to the pair can be broken down into these three points. For a new solution, the larger elements of the better solution and the smaller ones of the weaker one are combined to make a new solution. For a given set of part sizes, the better solution will always take precedence. Each pair of selected solutions generates a new solution. In this step, fresh solutions are not inserted into the matrix. In order to give them an opportunity to grow on the reef, they are kept in a list of new solutions (released into the water).

Internal reproduction operator:
This operator makes use of the solutions that were left behind after the external reproduction process was completed. It's this ratio, 1-Fb, which tells us how many times we've used an internal reproduction operator to solve a problem. Each answer is subject to a random mutation thanks to the internal reproduction operator. A whole new solution is created by reversing the existing solution's bits.

Setting new solutions:
In this step, the likelihood of each solution being placed in the matrix cells is evaluated. In this manner, a new solution is generated from a random cell in the matrix. It will be added to the population in that cell if the identification of this solution is found there. the new solution can be added to the cell. A different random cell should be used if the new solution cannot be placed in that cell. The number of times a solution can attempt to be placed in the matrix is determined by the user and can range from 1 to h.

Asexual reproduction operator:
In this case, the reef's best solutions are being re-created by this operator. The solutions to the matrix are ranked in decreasing order of their health function values for this purpose. Then, a portion of the sorted list's first members is chosen to be reproduced. It's possible to develop a new solution that's identical to the original one. As a result of the problem's user-specified parameter Fa, a specific percentage of the population will be selected for asexual reproduction. As a final step, the new solutions are inserted into the matrix in the same way as before.

Depredation:
This is accomplished by selecting a fraction Fd of the poorest solutions from the previously ordered list. Fd can also be set by the user. If the generated random number is less than the parameter Pd, then the related solution is removed from the matrix and its corresponding cell is released. This process is repeated for each of these selected solutions. Probability of depredation is the Pd parameter, which can be set by the user. The proportion of deprecated solutions should be modest in the early iterations. With many vacant cells left over, fresh solutions can be inserted and grown in the matrix. A bigger number of subpar solutions should be discarded after numerous iterations. 0.1/k is added to Pd after each iteration, starting with a value of 0. (k is the number of iterations). The values of Fa and Fd should be chosen in such a way that Fa + Fd 1 in order to avoid the operations of asexual reproduction and depredation being overlapping.
After the stated number of iterations of creating populations, CRO will be shut down. The CRO's ultimate solution is determined by the matrix's best possible solution. Using this model, the ideal distance in the Minkowski metric can be found.

Prediction
Each predicting module outputs a P-dimensional vector Y that represents the projected values based on the most recent weight values and ST attributes gleaned from the lower levels. Using linear stretching, the predicted values are transferred back to the original scale and the actual prediction values are produced. A feature set is created by taking the predicted values from each subsequent forecasting module and transforming them back into their original image size before using them as dummy values in the final or "pth" forecasting module. This ensures that the forecasting process will not stop until all of the predicted values have been obtained. Thus, data from all timestamps can be used without conflicting with ST analysis' causality constraint.

Dummy Image/Data Tuning

The predicted image (pth) is compared to the original picture (It) of the time stamp at which it was generated. Once again, the forecasting process inside each forecasting module begins when the forecasted image falls below a certain degree of accuracy, which is stated in terms of a threshold value (th). As a result, each forecasting module's anticipated images and the weight values within the KNN model are fine-tuned. There will be no change in the forecasted image until the threshold th is met or the final forecasting module stops improving. The correctness of the data/image adjustment was evaluated using the mean absolute error (MAE) in our suggested model. The dummy imagery is regarded to be the predicted/reconstructed imagery of the appropriate time-series sequence after the termination condition is met.

IV. RESULTS AND DISCUSSION

Experiments on anticipating missing photos in a sequence of derived remote sensing imagery are described in this section.

Dataset and Study Area

For eight consecutive years from 2004 to 2011, NDVI images were taken over West Bengal's Bardhhaman district, which is located in the state of West Bengal (see **Fig 2**) [48].

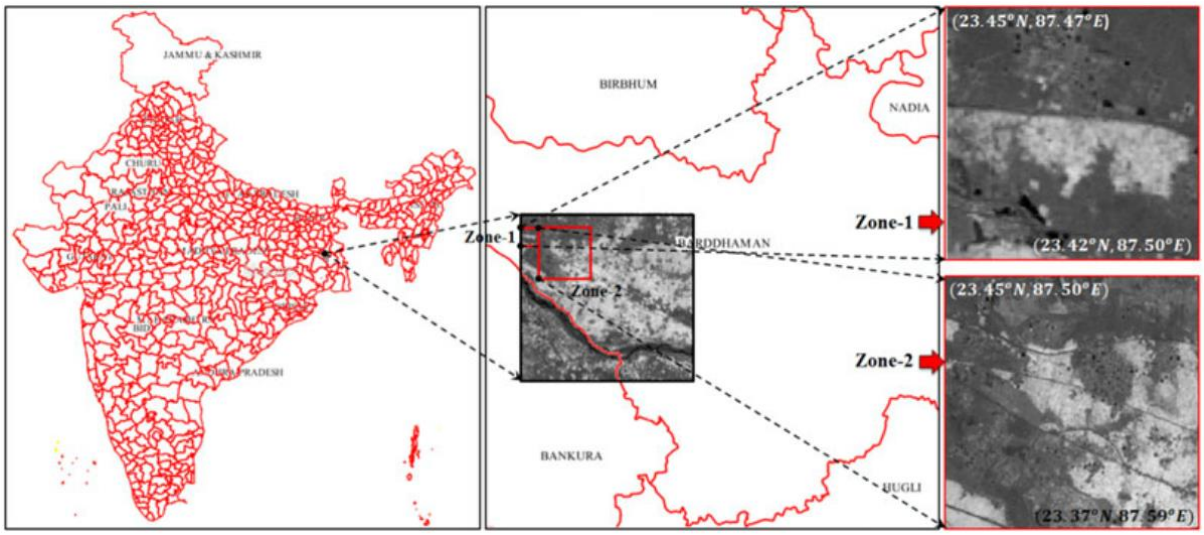

Fig 2. Study area West Bengal, India in the district of Bardhhaman.

It was then used to construct an NDVI raster from raw satellite imagery using ERDAS IMAGINE [49]. We've narrowed the scope of our study to two areas, each with thousands of pixels. **Table 1** summarises the zone-specific information. In addition, because the NDVI photos from 2007 and 2009 are absent, the empirical investigation was conducted twice. A random selection was made from among the years' worth of missing photos.

Table 1. Details of study regions in the barddhaman district

Zones	Bounding box details		
	Number of pixels	Top-Left	Bottom-Right
1	102 400	23.45 ◦N, 87.54 ◦E	23.37 ◦N, 87.60 ◦E
2	10 000	23.45 ◦N, 87.47 ◦E	23.42 ◦N, 87.50 ◦E

Evaluation Metrics

To determine how well a statistical or machine learning model is functioning, performance measurements are utilised. Each project must assess the models or techniques used in machine learning. There are many various assessment methods that can be used to test a model. These include classification performance, exponential loss, confusion matrix, and other metrics. By "accuracy," we typically refer to recognition rate, which is the ratio of the number of correctly predicted events to the total amount of input samples. Inaccurate categorization are punished using logarithmic loss, also referred to as log loss. The ability to compare the reconstructed data to the actual observed data allowed us to evaluate the missing locations. We first generated the simulated cloud-covered photos by removing random patches from the target images (with different forms, areas, and scattered placements). When we looked at all the pixels in the simulated cloud-covered regions, we compared them to the actual pixels to see how accurate our reconstructions were.

We were able to get the projected value (x_i) back for pixel I with an actual value of x_i in a particular band. $e_i = |x_i - \hat{x}_i|$ can be used to compute the reconstruction error. The greater the precision, the lower the value of e should be. To quantify cloud-covered regions' total reconstruction accuracy, we used a combination of the NMSE, MAE, MAE, and CC indices.

$$NMSE = \frac{1}{N}\sum_{i\in R}\frac{(x_i-\hat{x}_i)^2}{u(x)u(\hat{x})} \quad (3)$$

$$MAE = \frac{1}{N}\sum_{i\in R}|x_i-\hat{x}_i| \quad (4)$$

$$MAPE = \frac{1}{N}\sum_{I\in R}|x_i-\hat{x}_i|/x_i \quad (5)$$

$$CC = \frac{\sum_{i\in R}(x_i-u(x_i))(\hat{x}_i-u(\hat{x}_i))}{\sqrt{\sum_{i\in R}(x_i-u(x))^2\sum_{i\in R}(\hat{x}_i-u(\hat{x}))^2}} \quad (6)$$

Difference between observed and predicted discoveries is represented by NMSE, MAE depicts when the data are of different sizes It can be shown by CC that the observed and recovered results have similar texture. In these tests, lower NMSE, MAE, and MAPE values indicate better reconstruction and are hence preferred to higher ones. All portions of the dataset will be identical if CC is greater than 1.

An improvement ratio index (IR) is used to measure the improvement of a given method in comparison to other approaches.

$$IR = \frac{V_{sota}-V_{prop}}{V_{sota}} \ for\ NMSE, MAE\ and\ MAPE; IR = \frac{V_{prop}-V_{sota}}{V_{sota}}\ for\ CC \quad (7)$$

Comparative (state-of-the-art) approaches are evaluated using V_{sota}, while the suggested method is evaluated using V_{prop}.

The proposed KNN is optimized by coral reef and its performance is validated by using different optimization models such as Fruit fly (FFA), BFO and BAT algorithms in terms of various metrics by considering different channels such as red, green, blue and Nir. These experiments are provided in **Table 2** and graphical chart of proposed IR with various techniques is presented in **Fig 3.**

Table 2. Accuracy Assessment of Renovation Results in the Relative Experiments.

Channel	Parameter	FFA-KNN	CRO-KNN	BFO-KNN	BAT-KNN	IR (%)
Red	NMSE	1.8484	1.4821	2.1013	1.8850	19.82
	MAE	78.3571	49.2581	67.8327	69.8642	27.38
	MAPE	8.9738	8.1894	8.9293	8.7937	6.87
	CC	0.8423	0.9216	0.8951	0.8808	2.96
Green	NMSE	0.9171	0.6029	0.6831	0.7494	11.74
	MAE	42.5354	32.1560	38.4329	40.3169	16.33
	MAPE	5.7463	4.8841	5.4308	5.3901	9.39
	CC	0.8867	0.9572	0.9261	0.9217	3.36
Blue	NMSE	1.1063	0.9593	1.0961	1.0860	11.67
	MAE	40.0724	34.1562	39.1762	38.4256	11.11
	MAPE	6.9451	6.1283	6.8646	7.0539	10.73
	CC	0.8807	0.9346	0.8915	0.9032	3.48
Nir	NMSE	6.3309	5.7624	6.2370	7.6751	7.61
	MAE	154.7382	118.2696	176.4823	156.4461	23.57
	MAPE	16.5881	14.8089	16.7976	17.3284	10.73
	CC	0.8187	0.8537	0.7896	0.7961	4.285

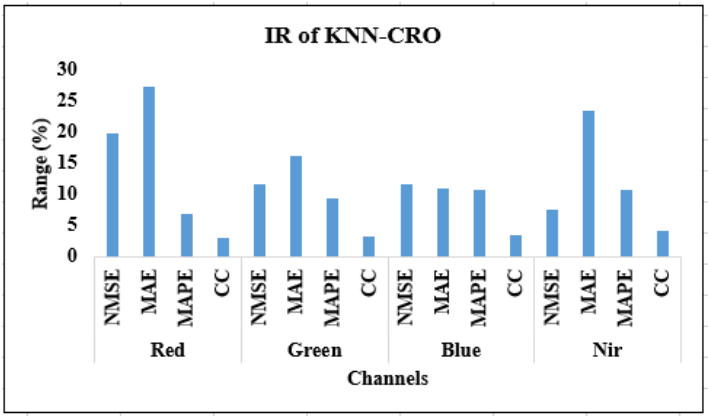

Fig 3. Graphical Chart of Proposed IR With Various Techniques in Terms of Various Metrics.

In the channel Red, the error of proposed model is very less than existing techniques, for instance, CRO has 49.25 MAE, FFA has 78.35, BFO has 67 and BAT has 69MAE. The improvement of proposed model with existing techniques in terms of MAE is 27.38. In C analysis, CRO has 92.16, FFA has 84.23, BFO has 89.51 and BAT has 88.08. In the channel of green for the analysis of MAPE, the CRO has less value and shows its better improvement, i.e 9.39%. Finally, the improvement of proposed model for NMSE is 11.74% and it has only 3.36% for CC analysis. In the analysis of MAPE for blue channel, the FFA achieved 6.94, BFO achieved 6.86, BAT achieved 7.05 and proposed model achieved 7.05. **Table 3** provides Indicating times in one epoch.

For NMSE, BFO achieved 1.0961, BAT achieved 1.086 and CRO achieved 0.9572. In the analysis of CC, FFA achieved 0.8807, BFO achieved 0.89, BAT achieved 0.90 and proposed model achieved 0.9346. When the channel is Nir, the proposed model achieved 5.76 for NMSE, 118.27 for MAE, 14.80 for MAPE and 0.85 for CC, where the BAT achieved 7.67 for NMSE, 156.44 for MAE, 17.32 for MAPE and 0.79 for CC and FFA achieved 6.33 for NMSE, 154.73 for MAE, 16.58 for MAPE and 0.81 for CC. **Fig 4** provides the comparison of proposed model in terms of MAE and CC. In the Red channel, the suggested model's error such as CRO (49.25 MAE), FFA (78.35), BFO (67), and BAT (69MAE). The suggested model outperforms previous methods by 27.38 in terms of MAE. Manhattan obtained 0.86 for the blue channel, 0.91 for the green channel, 0.86 for the red channel, and 0.79 for the Nir channel in the study of CC, while Euclidean obtained 0.88 for the blue channel, 0.90 for the green channel, 0.85.

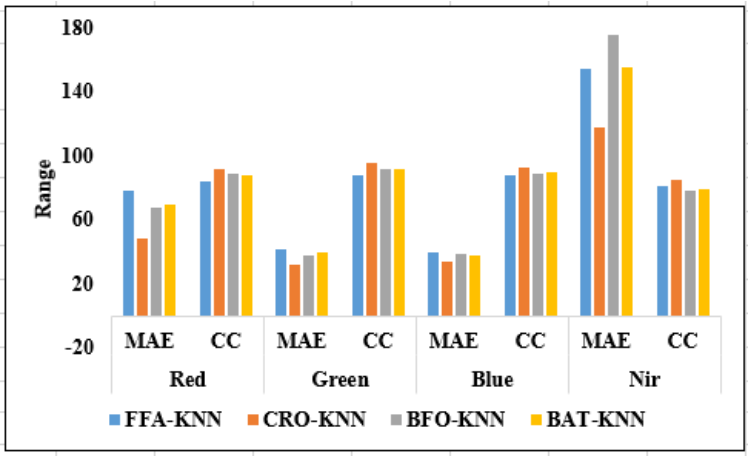

Fig 4. Graphical presentation of proposed model with various techniques in terms of CC and MAE.

Table 3. Computational Difficulty of the Four Reconstruction Approaches. * Indicating times in one epoch.

	BAT-KNN	BFO-KNN	FFA-KNN	CRO-KNN
Timing/s	378 + 78	1870*+582	297 + 116	570*+196
Memory/GB	0.3 + 0.4	2.1 + 2.5	0.5 + 0.8	1.8 + 2.1
GPU/GB	/	3.2 + 3.6	/	1.0 + 1.2

The existing techniques such as BAT, BFO and FFA achieved nearly 379s, 1870s and 297s, where the proposed model achieved 570s. In the analysis of memory, BAT has 0.3GB, BFO has 2.1GB, FFA has 0.5GB and proposed model has 1.8GB. From this analysis, it is clearly proving that the proposed model has better computational complexity than existing models. **Table 4** provides the validation analysis of KNN by changing the distance metrics, which is optimized by CRO.

Table 4. Accuracy Valuation of Reconstruction Results Produced From the Various Distance Metrics.

	Manhattan				**Euclidean**				**Minkowski**			
Metrics	NMSE	MAE	MAPE	CC	NMSE	MAE	MAPE	CC	NMSE	MAE	MAPE	CC
Blue	2.8026	36.3328	9.6273	0.8674	3.3876	37.3584	10.8450	0.8805	0.9771	34.1094	7.0427	0.9027
Green	1.1994	39.0973	7.1844	0.9148	1.4682	38.9068	7.6234	0.9076	0.5735	32.6492	5.1876	0.9481
Red	2.6458	51.0496	10.8735	0.8647	3.7065	53.2561	11.4183	0.8514	1.6908	46.7392	9.3291	0.9294
Nir	27.0126	149.8232	29.2901	0.7949	20.5649	168.5934	30.2819	0.7683	9.4307	115.9650	17.4318	0.8419

In the analysis of CC, Manhattan achieved 0.86 for blue channel, 0.91 for green channel, 0.86 for red channel and 0.79 for Nir channel, Euclidean achieved 0.88 for blue channel, 0.90 for green channel, 0.85 for red channel and 0.76 for Nir channel, where the proposed Minkowski achieved 0.90 for blue channel, 0.94 for green channel, 0.92 for red channel and 0.84 for Nir channel. When compared with all channels, Nir has high MAE for all distance metrics, for instance, Manhattan achieved 149.82, Euclidean achieved 168.59 and proposed model achieved 115.96. **Fig 5** shows the graphical representation of proposed distance metric model than existing technique.

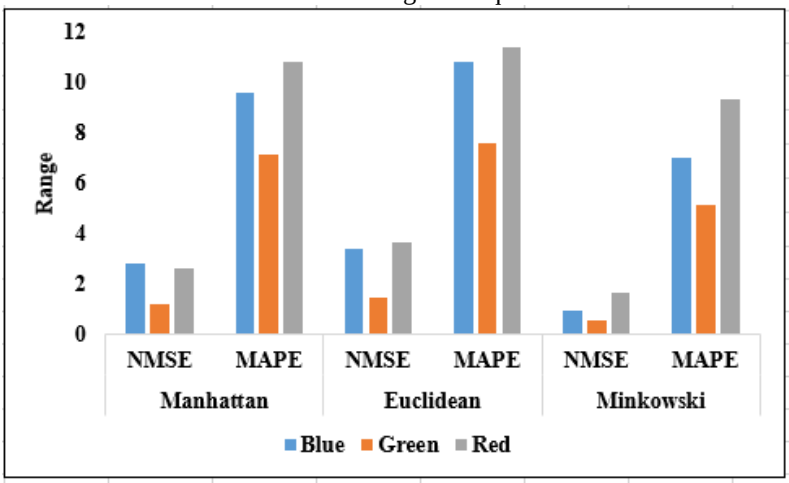

Fig 5. Graphical Representation of proposed distance metrics in terms of NMSE and MAPE.

In the NMSE analysis, Manhattan distance achieved nearly 1 to 2 for all channels and 27.01 for Nir channel, Euclidean distance achieved nearly 1 to 3 for red, green and blue channels and 20.56 for Nir channel and proposed Minkowki distance achieved 9.43 for Nir channel and 0.5 to 1 for all channels. Instead of the previous models' close to 9 to 11 for red, green, and blue channels and close to 29 to 30 for Nir channel, the proposed distance metric obtained reduced MAPE for all channels, such as nearly 7 to 9 for red, green, and blue channels and 17.43 for Nir channel. The method can be expanded to predict missing raw satellite photos that contain a variety of spectral bands and layers. **Fig 6** shows the graphical comparison of proposed distance of KNN model.

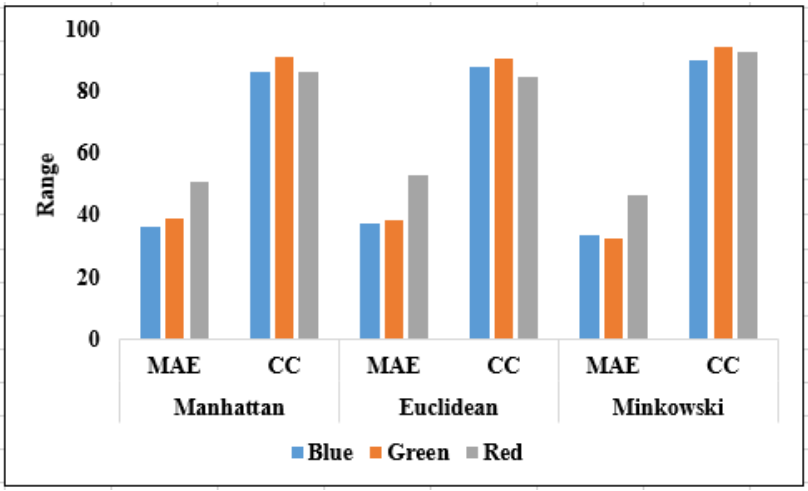

Fig 6. Graphical comparison of proposed distance of KNN model.

We were able to assess the missing places by contrasting the rebuilt data with the actual observed data. In order to create the fake cloudy shots, we first removed arbitrary regions from the target images. when there are disparate sizes of the data CC can demonstrate that the texture of the observed and retrieved findings is similar. Lower NMSE, MAE, and MAPE values are favoured than higher ones in these tests since they imply better reconstruction. Nir has strong MAE for all distance metrics compared to all channels, such as Manhattan attained 149.82, Euclidean attained 168.59, and suggested model attained 115.96.

V. CONCLUSION

Remote sensing analyses are frequently hampered by missing data, which could not be generated because source satellite imagery was unavailable. In order to anticipate missing data in remote sensing time series, this research suggests a weight optimised ML approach. Based on the Minkowski metric for KNN, the proposed prediction method is weight optimization. An advantage of this new strategy over previous ones is the ability to take advantage of remote sensing data from earlier

and later in the series without conflicting with the causality requirement. It's a great approach to get the most out of the knowledge you already have. Two spatial zones in Barddhaman, India, comprised of hundreds of pixels, have been studied to forecast missing NDVI imagery from 2004 to 2011. When compared to existing machine learning models, the suggested model outperforms the current state of the art ML-learning-based ST prediction models and conventional time series prediction methods. The proposed distance metric achieved less MAPE for all channels such as nearly 7 to 9 for red, green and blue channels and 17.43 for Nir channel, where the existing models achieved nearly 9 to 11 for red, green and blue channels, then nearly 29 to 30 for Nir channel. It is possible to extend the approach to forecast missing raw satellite images, which includes many bands/layers of spectral information. By leveraging deep learning models to run the forecasting modules in parallel, the suggested methodology can also handle very large remote sensing datasets.

Data Availability
No data was used to support this study.

Conflicts of Interests
The author(s) declare(s) that they have no conflicts of interest.

Funding
No funding agency is associated with this research.

Competing Interests
There are no competing interests.

References

[1]. Y. Cheng, K. Zhou, J. Wang, and J. Yan, “Big Earth Observation Data Integration in Remote Sensing Based on a Distributed Spatial Framework,” Remote Sensing, vol. 12, no. 6, p. 972, Mar. 2020, doi: 10.3390/rs12060972.

[2]. A. Sharma, K. Singh, and D. Koundal, “A novel fusion based convolutional neural network approach for classification of COVID-19 from chest X-ray images,” Biomedical Signal Processing and Control, vol. 77, p. 103778, Aug. 2022, doi: 10.1016/j.bspc.2022.103778.

[3]. A. G. Dekker et al., “Intercomparison of shallow water bathymetry, hydro-optics, and benthos mapping techniques in Australian and Caribbean coastal environments,” Limnology and Oceanography: Methods, vol. 9, no. 9, pp. 396–425, Sep. 2011, doi: 10.4319/lom.2011.9.396.

[4]. L. Jiao, W. Sun, G. Yang, G. Ren, and Y. Liu, “A Hierarchical Classification Framework of Satellite Multispectral/Hyperspectral Images for Mapping Coastal Wetlands,” Remote Sensing, vol. 11, no. 19, p. 2238, Sep. 2019, doi: 10.3390/rs11192238.

[5]. S. Khanna, M. Santos, S. Ustin, K. Shapiro, P. Haverkamp, and M. Lay, “Comparing the Potential of Multispectral and Hyperspectral Data for Monitoring Oil Spill Impact,” Sensors, vol. 18, no. 2, p. 558, Feb. 2018, doi: 10.3390/s18020558.

[6]. S. Kaur, S. Gupta, S. Singh, D. Koundal, and A. Zaguia, “Convolutional neural network based hurricane damage detection using satellite images,” Soft Computing, vol. 26, no. 16, pp. 7831–7845, Feb. 2022, doi: 10.1007/s00500-022-06805-6.

[7]. J. Schindelin, C. T. Rueden, M. C. Hiner, and K. W. Eliceiri, “The ImageJ ecosystem: An open platform for biomedical image analysis,” Molecular Reproduction and Development, vol. 82, no. 7–8, pp. 518–529, Jul. 2015, doi: 10.1002/mrd.22489.

[8]. D. A. Stow et al., “Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems,” Remote Sensing of Environment, vol. 89, no. 3, pp. 281–308, Feb. 2004, doi: 10.1016/j.rse.2003.10.018.

[9]. G. P. Asner, “Cloud cover in Landsat observations of the Brazilian Amazon,” International Journal of Remote Sensing, vol. 22, no. 18, pp. 3855–3862, Jan. 2001, doi: 10.1080/01431160010006926.

[10]. G. Hmimina et al., “Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground-based NDVI measurements,” Remote Sensing of Environment, vol. 132, pp. 145–158, May 2013, doi: 10.1016/j.rse.2013.01.010.

[11]. I. Garonna, R. de Jong, and M. E. Schaepman, “Variability and evolution of global land surface phenology over the past three decades (1982–2012),” Global Change Biology, vol. 22, no. 4, pp. 1456–1468, Feb. 2016, doi: 10.1111/gcb.13168.

[12]. M. A. WHITE et al., “Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006,” Global Change Biology, vol. 15, no. 10, pp. 2335–2359, Sep. 2009, doi: 10.1111/j.1365-2486.2009.01910.x.

[13]. J. v. Buttlar, J. Zscheischler, and M. D. Mahecha, “An extended approach for spatiotemporal gapfilling: dealing with large and systematic gaps in geoscientific datasets,” Nonlinear Processes in Geophysics, vol. 21, no. 1, pp. 203–215, Feb. 2014, doi: 10.5194/npg-21-203-2014.

[14]. M. K. Goyal, A. Sharma, and R. Y. Surampalli, “Remote Sensing and<scp>GIS</scp>Applications in Sustainability,” Sustainability, pp. 605–626, Mar. 2020, doi: 10.1002/9781119434016.ch28.
[15]. J. E. Patino and J. C. Duque, “A review of regional science applications of satellite remote sensing in urban settings,” Computers, Environment and Urban Systems, vol. 37, pp. 1–17, Jan. 2013, doi: 10.1016/j.compenvurbsys.2012.06.003.
[16]. M. He, Y. Hu, N. Chen, D. Wang, J. Huang, and K. Stamnes, “High cloud coverage over melted areas dominates the impact of clouds on the albedo feedback in the Arctic,” Scientific Reports, vol. 9, no. 1, Jul. 2019, doi: 10.1038/s41598-019-44155-w.
[17]. R. E. Wolfe, D. P. Roy, and E. Vermote, “MODIS land data storage, gridding, and compositing methodology: Level 2 grid,” IEEE Transactions on Geoscience and Remote Sensing, vol. 36, no. 4, pp. 1324–1338, Jul. 1998, doi: 10.1109/36.701082.
[18]. K. Wang, S. E. Franklin, X. Guo, and M. Cattet, “Remote Sensing of Ecology, Biodiversity and Conservation: A Review from the Perspective of Remote Sensing Specialists,” Sensors, vol. 10, no. 11, pp. 9647–9667, Nov. 2010, doi: 10.3390/s101109647.
[19]. T. J. Schmugge, W. P. Kustas, J. C. Ritchie, T. J. Jackson, and A. Rango, “Remote sensing in hydrology,” Advances in Water Resources, vol. 25, no. 8–12, pp. 1367–1385, Aug. 2002, doi: 10.1016/s0309-1708(02)00065-9.
[20]. C. J. Tomlinson, L. Chapman, J. E. Thornes, and C. Baker, “Remote sensing land surface temperature for meteorology and climatology: a review,” Meteorological Applications, vol. 18, no. 3, pp. 296–306, Aug. 2011, doi: 10.1002/met.287.
[21]. R. P. Gupta, “Remote Sensing Geology,” Springer: Berlin/Heidelberg, Germany, 2017.
[22]. Shirzadi, A.; Soliamani, K.; Habibnejhad, M.; Kavian, A.; Chapi, K.; Shahabi, H.; Chen, W.; Khosravi, K.; Pham, B.T. Shallow Landslide Susceptibility Mapping. Sensors 2018, 18, 3777.
[23]. N. Xu, “The Application of Deep Learning in Image Processing is Studied Based on the Reel Neural Network Model,” Journal of Physics: Conference Series, vol. 1881, no. 3, p. 032096, Apr. 2021, doi: 10.1088/1742-6596/1881/3/032096.
[24]. I. Leifer et al., “State of the art satellite and airborne marine oil spill remote sensing: Application to the BP Deepwater Horizon oil spill,” Remote Sensing of Environment, vol. 124, pp. 185–209, Sep. 2012, doi: 10.1016/j.rse.2012.03.024.
[25]. Z. Chen, B. Cong, Z. Hua, K. Cengiz, and M. Shabaz, “Application of clustering algorithm in complex landscape farmland synthetic aperture radar image segmentation,” Journal of Intelligent Systems, vol. 30, no. 1, pp. 1014–1025, Jan. 2021, doi: 10.1515/jisys-2021-0096.
[26]. M. Lauer and S. Aswani, “Integrating indigenous ecological knowledge and multi-spectral image classification for marine habitat mapping in Oceania,” Ocean & Coastal Management, vol. 51, no. 6, pp. 495–504, Jan. 2008, doi: 10.1016/j.ocecoaman.2008.04.006.
[27]. E. Adam, O. Mutanga, J. Odindi, and E. M. Abdel-Rahman, “Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers,” International Journal of Remote Sensing, vol. 35, no. 10, pp. 3440–3458, Apr. 2014, doi: 10.1080/01431161.2014.903435.
[28]. S. L. Bangare, G. Pradeepini, and S. T. Patil, “Neuroendoscopy Adapter Module Development for Better Brain Tumor Image Visualization,” International Journal of Electrical and Computer Engineering (IJECE), vol. 7, no. 6, p. 3643, Dec. 2017, doi: 10.11591/ijece.v7i6.pp3643-3654.
[29]. B. Chen, B. Huang, L. Chen, and B. Xu, “Spatially and Temporally Weighted Regression: A Novel Method to Produce Continuous Cloud-Free Landsat Imagery,” IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 1, pp. 27–37, Jan. 2017, doi: 10.1109/tgrs.2016.2580576.
[30]. S. K. Padhee and S. Dutta, “Spatio-Temporal Reconstruction of MODIS NDVI by Regional Land Surface Phenology and Harmonic Analysis of Time-Series,” GIScience & Remote Sensing, vol. 56, no. 8, pp. 1261–1288, Aug. 2019, doi: 10.1080/15481603.2019.1646977.
[31]. P. Rani, R. Kumar, and A. Jain, “Multistage Model for Accurate Prediction of Missing Values Using Imputation Methods in Heart Disease Dataset,” Lecture Notes on Data Engineering and Communications Technologies, pp. 637–653, 2021, doi: 10.1007/978-981-15-9651-3_53.
[32]. A. Elhassan, S. M. Abu-Soud, F. Alghanim, and W. Salameh, “ILA4: Overcoming missing values in machine learning datasets – An inductive learning approach,” Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 7, pp. 4284–4295, Jul. 2022, doi: 10.1016/j.jksuci.2021.02.011.
[33]. E. Adam, O. Mutanga, J. Odindi, and E. M. Abdel-Rahman, “Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers,” International Journal of Remote Sensing, vol. 35, no. 10, pp. 3440–3458, Apr. 2014, doi: 10.1080/01431161.2014.903435.
[34]. A. Singh, S. Kushwaha, M. Alarfaj, and M. Singh, “Comprehensive Overview of Backpropagation Algorithm for Digital Image Denoising,” Electronics, vol. 11, no. 10, p. 1590, May 2022, doi: 10.3390/electronics11101590.

[35]. D. J. Park, M. W. Park, H. Lee, Y.-J. Kim, Y. Kim, and Y. H. Park, "Development of machine learning model for diagnostic disease prediction based on laboratory tests," Scientific Reports, vol. 11, no. 1, Apr. 2021, doi: 10.1038/s41598-021-87171-5.

[36]. L. Zhu, G. Wang, F. Huang, Y. Li, W. Chen, and H. Hong, "Landslide Susceptibility Prediction Using Sparse Feature Extraction and Machine Learning Models Based on GIS and Remote Sensing," IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, 2022, doi: 10.1109/lgrs.2021.3054029.

[37]. R. Sedona, G. Cavallaro, J. Jitsev, A. Strube, M. Riedel, and J. Benediktsson, "Remote Sensing Big Data Classification with High Performance Distributed Deep Learning," Remote Sensing, vol. 11, no. 24, p. 3056, Dec. 2019, doi: 10.3390/rs11243056.

[38]. M. Gangappa, C. Kiran, and P. Sammulal, "Techniques for Machine Learning based Spatial Data Analysis: Research Directions," International Journal of Computer Applications, vol. 170, no. 1, pp. 9–13, Jul. 2017, doi: 10.5120/ijca2017914643.

[39]. J. N. Goetz, A. Brenning, H. Petschko, and P. Leopold, "Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling," Computers & Geosciences, vol. 81, pp. 1–11, Aug. 2015, doi: 10.1016/j.cageo.2015.04.007.

[40]. A. M. Youssef, H. R. Pourghasemi, Z. S. Pourtaghi, and M. M. Al-Katheeri, "Erratum to: Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia," Landslides, vol. 13, no. 5, pp. 1315–1318, Dec. 2015, doi: 10.1007/s10346-015-0667-1.

[41]. H. Wang, A. Sharma, and M. Shabaz, "Research on digital media animation control technology based on recurrent neural network using speech technology," International Journal of System Assurance Engineering and Management, vol. 13, no. S1, pp. 564–575, Mar. 2022, doi: 10.1007/s13198-021-01540-x.

[42]. A. Mehbodniya, J. L. Webber, M. Shabaz, H. Mohafez, and K. Yadav, "RETRACTED ARTICLE: Machine Learning Technique to Detect Sybil Attack on IoT Based Sensor Network," IETE Journal of Research, vol. 69, no. 10, Dec. 2021, doi: 10.1080/03772063.2021.2000509.

[43]. B. James and B. Yoshua, "Random Search for Hyper-Parameter Optimization," J. Mach. Learn. Res. 13 (1) (2012) 281305.

[44]. A. Moubayed, M. Injadat, A. Shami, and H. Lutfiyya, "DNS Typo-Squatting Domain Detection: A Data Analytics & Machine Learning Based Approach," 2018 IEEE Global Communications Conference (GLOBECOM), Dec. 2018, doi: 10.1109/glocom.2018.8647679.

[45]. C. Gambella, B. Ghaddar, and J. Naoum-Sawaya, "Optimization problems for machine learning: A survey," European Journal of Operational Research, vol. 290, no. 3, pp. 807–828, May 2021, doi: 10.1016/j.ejor.2020.08.045.

[46]. J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy K-nearest neighbor algorithm," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-15, no. 4, pp. 580–585, Jul. 1985, doi: 10.1109/tsmc.1985.6313426.

[47]. W. Zuo, D. Zhang, and K. Wang, "On kernel difference-weighted k-nearest neighbor classification," Pattern Analysis and Applications, vol. 11, no. 3–4, pp. 247–257, Jan. 2008, doi: 10.1007/s10044-007-0100-z.

[48]. "USGS EarthExplorer: Land Processes Distributed Active Archive Center," 2014. [Online]. Available: https://lpdaac.usgs.gov/data_access/usgs_earthexplorer

[49]. "ERDAS IMAGINE: Hexagon Geospatial," 2014. [Online]. Available: http://www.hexagongeospatial.com/products/remote-sensing/erdasimagine/overview