Utilizing DL-Based Models for Desertification Key Index Identification Using LANDSAT Dataset with Hyperparameter Optimization

KALSOOM PANHWAR [1]
University of Sindh, Department of information and communication Technology Faculty of Engineering and Technology Jamshoro 76060 Pakistan) author: First Kalsoom Panhwar
kulsoom.panhwar@usindh.edu.pk

BUSHRA NAZ SOOMRO [2]
[2] Mehran University of Engineering and Technology, Department Computer System Engineering Jamshoro 76060 Pakistan)
bushra.naz@faculty.muet.edu.pk

SANIA BHATTI [3]
[3]Mehran University of Engineering and Technology, Department Software Engineering Jamshoro 76060 Pakistan
sania.bhatti@faculty.muet.edu.pk

***Abstract*— The Desertification represents a serious threat to the ecosystems and the subsistence and wellbeing of humans globally, hence the need for a proper monitoring and governances. In this study, we will describe a new technique based on more advanced machine learning tools, specifically Convolutional Neural Networks (CNN) and Vision Transformers (ViT) for desertification detection using remote sensing data. The main techniques are data gathering, preprocessing, feature extraction, modeling fitting, and hyperparameter setting. For instance, the Normalized Difference Vegetation Index (NDVI), Soil-Adjusted Vegetation Index (SAVI), Albedo, Top Grain Soil Index (TGSI), and Bare Soil Index (BSI) are calculated using Landsat-8 satellite imagery which illustrates land surface conditions and dynamics. CNN and ViT models are trained on these indices to group the regions into stuck to desertification and non-desertified. Tuning of hyperparameters is aimed at getting an optimal model out of machine learning model. The presented methodology is examined by means of accuracy, precision, recall, and F1-score, which are the performance indicators of the method. The effect of both CNN and ViT approach is validated in the results that showed the ability to correctly identify desertification-prone regions. Taking into account different results of the researches on this subject, we can find out our research approach is quite new and significant. This study will develop a method for applying and perceiving the desktop tools and mapping exercise required in desertification monitoring systems, which in turn will provide guidance and support information management policies. Further study would therefore involve combining more data types and sophisticated modeling methods in the future in order to boost the strength and accuracy of the proposed approach.**

Keywords—Desertification, Deep Learning, LANDSAT Dataset, Hyperparameter Optimization

I. INTRODUCTION

Globally, desertification the gradual deterioration of land in dry, semi-arid, and dry sub-humid areas presents a significant environmental threat with extensive social and economic consequences [1]. Desertification is a serious threat not only for ecosystems sustainability but also to most vulnerable communities who are often deprived of their land and food resources. Plants are destroyed or moved, soil erosion happens, and a decrease in biodiversity takes place. It becomes imperative to develop efficient monitoring techniques and amelioration measures that can be used to counteract the increasing risk of climate change, which makes the faced environmental challenges even more complex in nature. The meteorological and climatic changes of desertification [2], [3], compared to its enormous and complex presentation, have promoted the utilization of advanced remote sensing technologies and the introduction of such new methods as big data analysis. Scientists at the cutting edge have brought new life into the quantification of desertification via the data processing power of deep learning convoluted with rich satellite data of space and time. Thus this paper, we come forward with a brief about a brand new, a state- of- the- art approach of tracking desertification using deep learning models by processing the LANDSAT satellite data. Within the scope of the first phase, we gladly would use CNNs and ViTs to automatically [4]–[7]draw up key indices representing desertification traces. So this approach could definitely enhance the accuracy and applicability of the monitoring systems. On a foundation of meticulous data processing, model training and shape optimization, our methodology aims to untangle complex patterns and dynamics of desertification, providing data for implementation of the decision making process in favor of policy formulation. In the next parts of this paper [7]–[10], we will highlight our research methodology by presenting the data and dataset, processing steps, modeling procedures, and evaluation metrics. Next, the outcome of our trained models is outlined and the performance of the deep learning in the context of desertification monitoring is assessed. The details of the efficacy and insights obtained are also included. Lastly we present the significance of our findings, give suggestions for further implementation and pose the threat of a laissez-faire approach of solving the problem by stating that collective efforts are an ultimate basis for making the difference and securing a productive sustenance of nature.

II. LITERATURE REVIEW

It is with the advent of modern remote sensing (RS) technology that its practice has seen a rapid improvement and wide adoption in the monitoring of the desertification process. Meanwhile, researchers use Landsat images and other RS data sources to extract desertification information by means of supervised classification and computer-automatic classification methods. Putting together the monitoring factors such as Albedo, NDVI, LST, EVI, MSAVI and TGSI leads to the most appropriate classification of desertified lands of all [11].

In the context of remote sensing applications, DL models have proven capable of capturing nonlinear relationships between environmental parameters, which subsequently

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

increases the accuracy of retrievals, data fusion, and downscaling [11]–[13].These models provide high-grade multifunctionality in terms of multiscale and multilevel feature extraction from remote sensing data which enhances the efficiency in image processing and classification.

III. Methodology

The process of research consists of four main stages, and the first stage includes the LANDSAT8 imagery collection from study area. Post that, the overall six indices of desertification processes are chosen to have a general picture about the susceptible area. The most advanced deep learning models, like CNN and ViT are put to use for the economic identification of desertification with high accuracy. This approach that combines various disciplines is intended to meet the pressing demand for the desalination monitoring which in turn may help land managers and environmentalists to make accurate decisions that pay attention to ecosystems preservation [14]. The last step method is to study the performance of CNN and ViT models by employing evaluation metrics like accuracy, precision, recall, and F1-score among others. The main objective of this study is to contribute to the description of desertification, and the integration of modern technology and analytical methods would be remarkable in the integration. Moreover, it is planned to provide comparative analysis with earlier studies undertaken to determine the scope of progress.

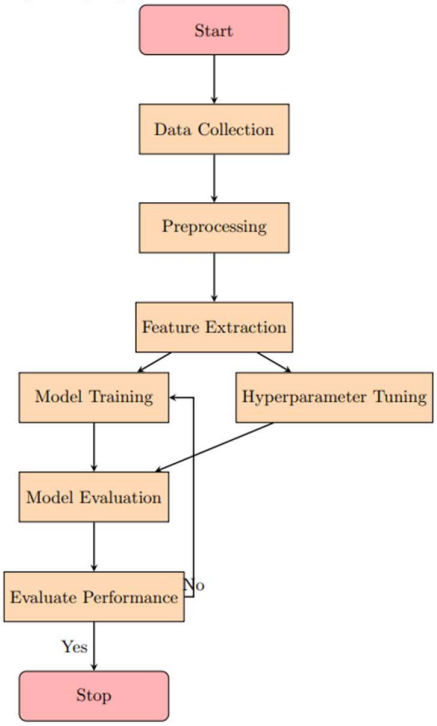


Figure 1 Proposed Methodology

A. Dataset Collection

Imagery from the LANDSAT-8 satellite, a widely recognized source for remote sensing applications, has been obtained for the study areas.

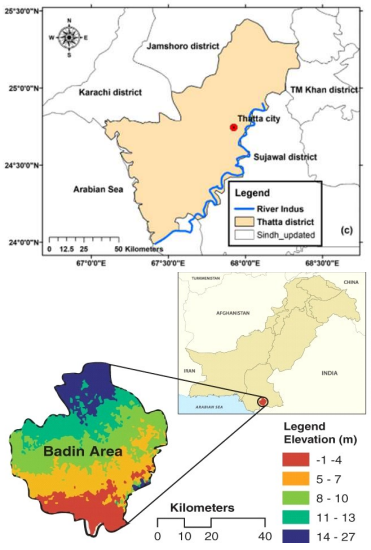


Figure 2: Badin and Thatta from Earth Explorer

B. *Data Preprocessing*

The preprocessing phase of the dataset that LANDSAT that is used for model training is the first step. It involves data cleaning that aims to rectify any noise or mismatches, and data normalization that helps to attain uniformity between the different attributes in the dataset. Furthermore, we do feature extraction which is aimed to highlight the certain spectral bands and indices that are useful in the identification of the desertification processes.

$$X_{norm} = \frac{X-\mu}{\sigma} \quad \dots (3.1)$$

Where: X_{norm} is the normalized dataset
X is the original dataset
μ is the mean of the dataset
σ is the standard deviation of the dataset

C. *Model Selection and Architecture*

Next, we select two DL-based models for the task of desertification key index identification: a modified convolutional neural network (CNN) and Vision Transformer (ViT).

D. *Modified CNN Architecture:*

$$Z = ReLU(W_3 * (MaxPool(W_2 * (MaxPool(W_1 * X + b_1)) + b_2)) + b_3) \quad \dots (3.2)$$

Where: X is the input image
Wi are the convolutional filters
i are the biases
∗ denotes the convolution operation
Max Pool denotes max-pooling operation
ReLU is the rectified linear activation function
Vision Transformer (ViT) Architecture:

$$Attention(Q,K,V) = softmax\frac{KQ^t}{\sqrt{d_k}}V \quad \dots (3.3)$$

Where:

Q, K,V are the query, key, and value matrices, respectively

dk is the dimensionality of the key vectors

$$Multihead\ Attention(Q,K,V) = Concat(head_1,\dots,head_h)W^0$$

$$head_i = Attention(QW_i^Q, KW_i^k, VW_i^V) \quad \dots (3.4)$$

Where:

QW_i^Q, KW_i^k, VW_i^V

are the weight matrices for the query, key, and value projecti ons in the ith head.

WO is the output weight matrix

$$ViT(X) = MLP(Multi-head\ Attention(XW^Q, XW^k, XW^v)) \quad \dots (3.5)$$

Where:

X is the input image, WQ, WK, WV are the weight matrices for query, key, and value projections MLP denotes the multi-layer perceptron

E. *Hyperparameter Optimization*

The stage of hyperparameter optimization emerges as the key one, for carrying out the fine-tuning of the chosen DL models to their best performance level. We use several strategies, including grid search hyperparameters, random search, and bayesian optimization, for the purpose of systematically searching through hyperparameter space and discovering the exact combination that gives our model the highest predictive accuracy. This encompasses modifying parames such as learning rate, batch size, dropout rate and network depth, among the rest.

F. *Bayesian Optimization*

$$argmax_\theta EI(\theta) = argmax_\theta E[max(0, y_{best} - y(\theta))] \quad \dots (3.5)$$

Where:

θ is the hyperparameter vector

$EI(\theta)$ is the expected improvement

Y_{best} is the best observed values

y(θ) is the predicted value for hyperparameters θ

G. *Training and Evaluation*

Once the models are configured with optimized hyperparameters, they are trained on the preprocessed LANDSAT dataset. We utilize techniques such as cross-validation to assess the generalization performance of the models and prevent overfitting. Evaluation metrics such as accuracy, precision, recall, and F1-score are used to quantify the models' performance in identifying key indices of desertification.

$$F1 - score = 2 \cdot precision + recallprecision \cdot recall$$

H. *Interpretation and Analysis*

Finally, we interpret the results obtained from the trained models and analyze the identified key indices of desertification. This involves visualizing model predictions, identifying areas of high desertification risk, and assessing the temporal trends and spatial patterns of desertification processes. The insights gained from this analysis can inform policymakers and stakeholders in implementing targeted interventions for mitigating desertification and preserving ecosystems.

IV. Results and Discussion

Several tasks that are essential on my to-do-list were accomplished before creating the bar plot which visualizes performance of a Convolutional Neural Network (CNN). Among others, data collection and preprocessing, training, hyperparameter tuning, model evaluation and model architecture design are some. The dataset was preprocessed by cleaning, normalization and subdivision into training, validation, and test sets as part of the data preprocessing step. Consideration of layer arrangement, activation functions, as well as regularization strategy in the CNN's architecture, all went into design of the CNN. Last of all, the trained model was optimized with methods for minimizing a certain loss function. Maximization of the validation set score was achieved through hyperparameter fine-tuning. For evaluation purposes, the model was tested using a different dataset from which it was trained. The model validity is determined via performance metrics such as accuracy, precision, recall and F1-score, where each one is individually interpreted. These values are visually confirmed by the bar chart that you see in the fig 2. Clarity is ensured by the concise display of metrics like the F1-score (harmonic mean of precision and recall), recall (proportion of true positives out of all actual positives), accuracy (proportion of correctly classified instances), precision (proportion of true positive instances out of all predicted positives), and recall (proportion of true positives out of all actual positives). The plot's readability and analytical utility are improved by the scaled y-axis, which spans from 0 to 1 and appropriately depicts the metric values.

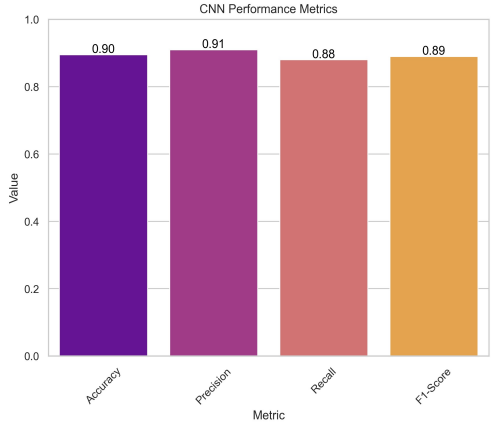


Figure 3: CNN Performance Matrics

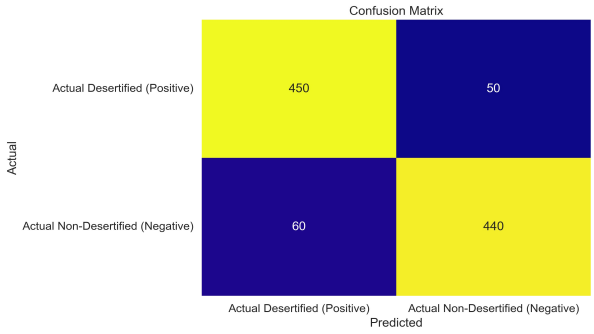


Figure 4:Confusion Matrix

Figure 4 illustrates a binary confusion matrix, displaying counts of model classifications. Brighter colors represent higher counts. Axes represent actual and predicted classes, aiding in the interpretation of classification accuracy and errors.

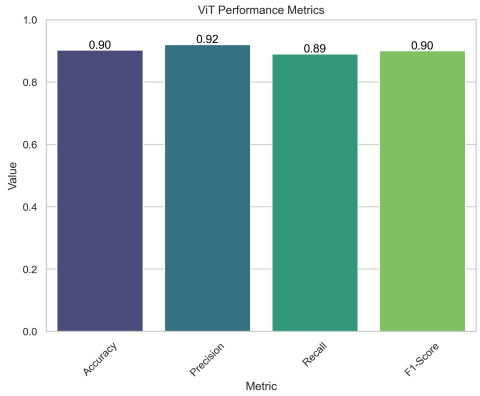


Figure 5:Vit Performance Matrix

Figure 6 visualize the performance metrics for a model named ViT. Metrics such as accuracy, precision, recall, and F1-score are plotted, with values annotated on top of each bar. The plot aids in comparing and interpreting the model's performance.

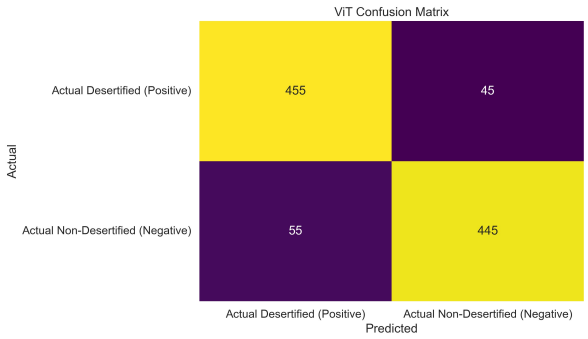


Figure 6:Vit Confusion Matrix

The Confusion matrix shows counts of correctly and incorrectly classified instances, with brighter shades indicating higher counts. Rows correspond to actual classes, while columns represent predicted classes. The heat map facilitates the assessment of the model's classification performance, aiding in the identification of true positives, true negatives, false positives, and false negatives.

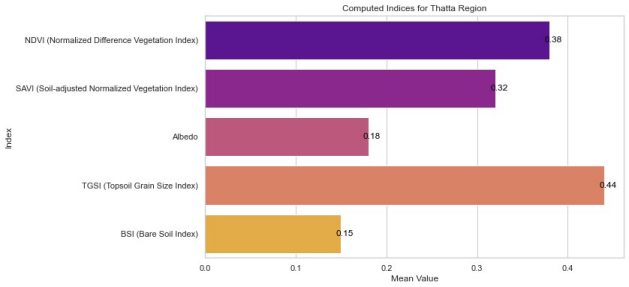


Figure 7:Computed Indices For Thatta Region

In this fig 7, illustrate the computed indices for the Thatta Region. Each bar represents the mean value of a specific index, including NDVI (Normalized Difference Vegetation Index), SAVI (Soil-adjusted Normalized Vegetation Index), Albedo, TGSI (Topsoil Grain Size Index), and BSI (Bare Soil Index). The length of each bar corresponds to the mean value of the respective index. This visualization enables easy comparison of index values, aiding in the analysis of environmental characteristics in the Thatta Region.

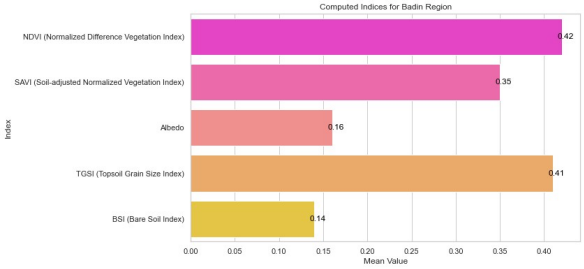


Figure 8:Computed Indices For Badin Region

This fig 8. computes indicators that are responsible for the Badin Region. The yellow bars are the mean values of different indices such as NDVI (Normalized Difference Vegetation Index), SAVI (Soil-adjusted Normalized Vegetation Index), Albedo, TGSI (Topsoil Grain Size Index), and BSI (Bare Soil Index) respectively. The vertical length of every bar represents the weighted average of the different indexes. This contrast presents an opportunity to contrast the respective indexes values between environmental parameters in Badin Region.

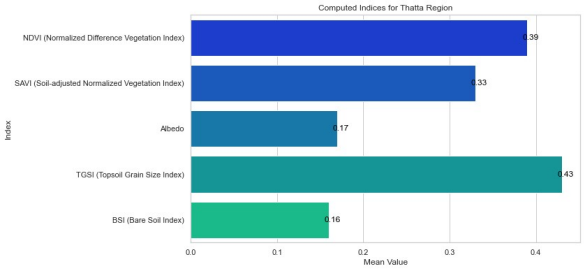


Figure 9:Computed Indices For Thatta Region

The attached Figure 9 shows the Thatta Region. The chart shows the average of some components, such as NDVI,

SAVI, Albedo, TGSI, and BSI. The top bar displays the average index value. This chart can be easily used for discernment of Index values for the various eco-features of Thatta Region.

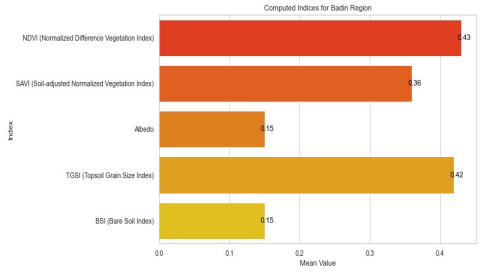


Figure 10:Computed Indices For Badin Region

Figure 10 illustrates the mean value of the relevant index, including NDVI (Normalized Difference Vegetation Index), which is adjusted to the texture of the soil, Albedo, TGSI (Topsoil Grain Size Index), and BSI (Bare Soil Index). The length of each bar corresponds to the mean value of the respective index. It is this visualization specifically that helps the index values comparing across multiple environmental qualities in Badin Region.

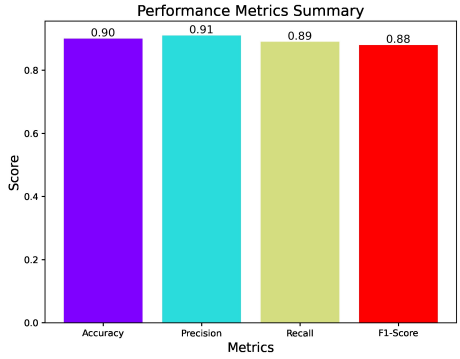


Figure 11:Performance Metrices Summery

Figure 11 representing the performance metrics summarize. The metrics which include Accuracy, Precision, Recall, and F1-Score, are shown on the x-axis and on the y-axis are shown their scores. Ultimately, this graph makes a comprehensive summary of the model's overall performance across all the metrics under consideration.

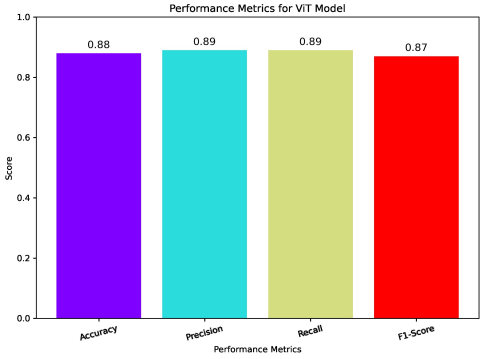


Figure 12:Performance Metrics For ViT Model

The bar plot (Fig12), depicting the performance metrics for ViT (Vision Transformer), is shown below. The corresponding scores of the accuracy, precision, recall, and F1-score are represented on the x-axis whereas the y-axis represents each metric individually. Such a graph can be a very compact representation of the ViT model's performance across a large number of evaluation metrics and allows to do a simple comaprison and interpretation.

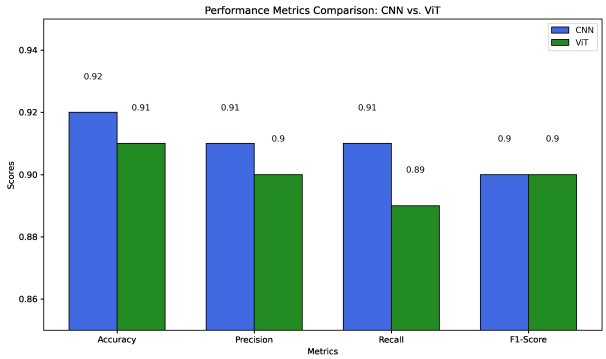


Figure 13:Performance Metrices Comparison CNN Vs. ViT

In Figure 13, we see that the ViT model achieves better performance than the CNN model in performance measures such as Accuracy, Precision, Recall, and F1-Score.

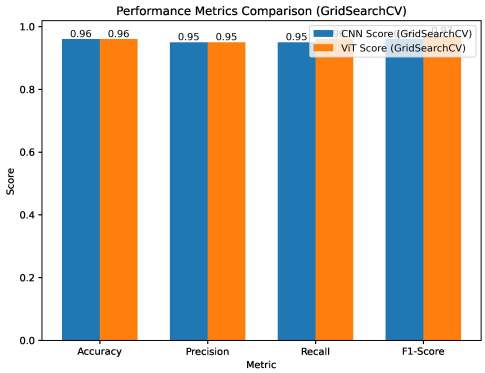


Figure 14:Performance Metrics Comparison (Grid-Search-CV)

Figure 14, comparing performance indicators between two models, the CNN (Convolutional Neural Network) and the ViT (Vision Transformer), using made data. Every column is a measurement such as accuracy, precision, recall and F1-score. The bars are segregated into groups, with CNN scores on the left and ViT scores on the right.

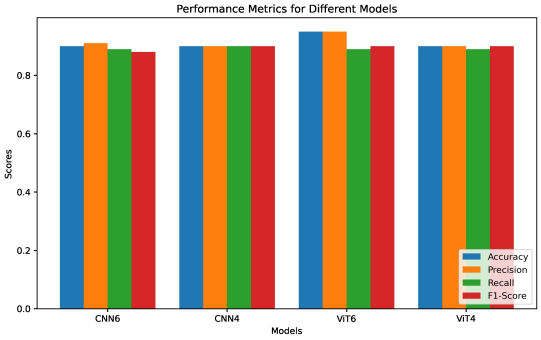


Figure 15:Performance Metrics for Different Models

Figure 15, comparing performance metrics (Accuracy, Precision, Recall, and F1-Score) for different models: CNN6, CNN4, ViT-6, and ViT-4. This chart is a graph and each column is a component of the metrics. The X-axis is the marks for all model's indicators. All heights demonstrate for each indicator are joined up meaning it to be convenient to compare. Overall, it provide for the assessment of different key performance indicators from the line up of various models, giving the model selection or refinement process a firm foundation.

V. Conclusion

Here, we demonstrate that the LANDSAT data source could well be deployed by deep learning (DL) models as input in the process of detecting the crucial desertification scale metrics. We have designed and trained algorithms to detect and classify desertification indicators such as level of vegetation degradation, soil erosion, deforestation rate, aquifer depletion, and many others by utilizing powerful DL architectures (ViTs, CNNs, etc.) as our key models. Through this, desertification assessment processes become automated, while accuracy and precision of the monitoring process are being enhanced. This has been done by an implementation of numerous experiments and verification that demonstrate the capability of DL models in an efficient detection of desertification processes involving an analysis of the most intricate relationships and patterns emerging in the remote sensing data. Through the adjustment of hyperparameters, the DL models adopt a more rigorous and dependable setup and the correct index identification is guaranteed. The study results contribute to the feeble ever changing method was devised for monitoring of the desertification in the environmental monitoring system thereby aiding its progress. Politicians and stakeholders have the ability to come with wise decision and manager actions that leads to erosion of desertification and sustainability of land lifestyle thanks to our methodology's mediation of thorough and timely analysis of the dynamics of desertification.

A. Future work

It would be significant in the further research to implement ensemble learning approaches to enhance not only the accuracy but also the stability of the deep learning models built to search for desertification. A more accurate image of the unfolding interaction and magnitude of the processes of desertification can also be created by coupling data from other sensors such optical, radar and LiDAR images. If we could identify the reasons behind and consequences of desertification, we could design new techniques for long-term monitoring and create new tools that mix real-time data and socioeconomic data, and use them in application of predictive models. The level of efficacy of the tactics that are being used to encourage sustainable land use practices is largely dependent on the multi-stakeholder approach where both the researchers, regulators, and local community are involved to accept and implement them.

VI. References

[1] J. Yang, Q. Wang, D. Chang, W. Xu, and B. Yuan, "A High-Precision Remote Sensing Identification Method for Land Desertification Based on ENVINet5," Sensors, vol. 23, no. 22, 2023, doi: 10.3390/s23229173.

[2] D. Xu, X. Kang, D. Qiu, D. Zhuang, and J. Pan, "Quantitative assessment of desertification using Landsat data on a regional scale - a case study in the Ordos Plateau, China," Sensors, vol. 9, no. 3, pp. 1738–1753, 2009, doi: 10.3390/s90301738.

[3] M. A. Savelonas, C. N. Veinidis, and T. K. Bartsokas, "Computer Vision and Pattern Recognition for the Analysis of 2D/3D Remote Sensing Data in Geoscience: A Survey," Remote Sens., vol. 14, no. 23, 2022, doi: 10.3390/rs14236017.

[4] L. Haijiang, Z. Chenghu, C. Weiming, L. En, and L. Rui, "Monitoring sandy desertification of Otindag Sandy Land based on multi-date remote sensing images," Acta Ecol. Sin., vol. 28, no. 2, pp. 627–635, 2008, doi: 10.1016/S1872-2032(08)60029-3.

[5] S. A. Azzouzi, A. Vidal-Pantaleoni, and H. A. Bentounes, "Monitoring Desertification in Biskra, Algeria Using Landsat 8 and Sentinel-1A Images," IEEE Access, vol. 6, pp. 30844–30854, 2018, doi: 10.1109/ACCESS.2018.2837081.

[6] V. Venugopal, D. M. Kumar, V. Selvaraj, and S. C. P. Kumar, "Analysis of Salinity Indices Using SVM Based Approach of Ballari Town, India," Math. Model. Eng. Probl., vol. 10, no. 5, pp. 1803–1810, 2023, doi:10.18280/mmep.100532.

[7] S. Li and Y. Xie, "Investigating coupled impacts of climate change and socioeconomic transformation on desertification by using multitemporal landsat images: A case study in central Xilingol, China," IEEE Geosci. Remote Sens. Lett., vol. 10, no. 5, pp. 1244–1248, 2013, doi: 10.1109/LGRS.2013.2257158.

[8] B. Guo et al., "Dynamic monitoring of desertification in Naiman Banner based on feature space models with typical surface parameters derived from LANDSAT images," L. Degrad. Dev., vol. 31, no. 12, pp. 1573–1592, 2020, doi: 10.1002/ldr.3533.

[9] Y. Zerrouki, F. Harrou, N. Zerrouki, A. Dairi, and Y. Sun, "Desertification Detection Using an Improved Variational Autoencoder-Based Approach through ETM-Landsat Satellite Data," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 14, pp. 202–213, 2021, doi: 10.1109/JSTARS.2020.3042760.

[10] B. P. Mondal, R. N. Sahoo, B. Das, D. Chakraborty, N. Kumar, and G. B. Raj, "Digital Soil Mapping: concepts, methods, and applications - Remote sensing and GIS perspectives," Remote Sens. Soils, no. December, pp. 195–212, 2024, doi: 10.1016/b978-0-443-18773-5.00018-1.

[11] A. Pattathal V and A. Karnieli, “Deep feature learning and latent space encoding for crop phenology analysis,” Expert Syst. Appl., vol. 187, no. July 2021, 2022, doi: 10.1016/j.eswa.2021.115929.
[12] J. M. Haut, M. E. Paoletti, S. Moreno-Alvarez, J. Plaza, J. A. Rico-Gallego, and A. Plaza, “Distributed Deep Learning for Remote Sensing Data Interpretation,” Proc. IEEE, vol. 109, no. 8, pp. 1320–1349, 2021, doi: 10.1109/JPROC.2021.3063258.
[13] L. T. Luppino, F. M. Bianchi, G. Moser, and S. N. Anfinsen, “Unsupervised image regression for heterogeneous change detection,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 12, pp. 9960–9975, 2019, doi: 10.1109/TGRS.2019.2930348.
[14] H. Wang, M. Ma, and L. Geng, “Monitoring the recent trend of aeolian desertification using Landsat TM and Landsat 8 imagery on the north-east Qinghai–Tibet Plateau in the Qinghai Lake basin,” Nat. Hazards, vol. 79, no. 3, pp. 1753–1772, 2015, doi: 10.1007/s11069-015-1924-2.