Design of A Water and Soil Classification Model for Satellite Images Via The Machine Learning Process

Arti Jane, Nandini Deshpande, Achal Dhorkhande, Yashashvi Dhoke, Chanchal Gothwad, Prachi Gajbhiye
CSE Department
Jhulelal Institute Of Technology, Nagpur-440001, Maharashtra
ndeshpande2001@gmail.com, jitcse.ndeshpande2001@gmail.com, jitcse.achaldorkhande02@gmail.com,
jitcse.yashashvidoke04@gmail.com, chanchalgothwad@gmail.com, prachigajbhiye@gmail.com

***Abstract*—The transmission of a variety of information about the target location in satellite photos is made possible by panchromatic and multispectral layers (ROI). Researchers will find it easier to identify related and distinctive site aspects thanks to these layers, which each highlight a particular part of the location. The enormous range of crop varieties, land cover types, soil types, and other associated features found in these locations may be identified using a variety of machine-learning algorithms. Each strategy differs from the others in terms of how effectively it works, how widely it can be applied, how challenging it is to put into practice, and how much money it will cost. In this study, we provide a brand-new deep-learning model for the on-the-fly, real-time region-based classification of satellite pictures. This model aims to advance upon the previous one. The large-scale temporal datasets for a particular region were first obtained using Google Earth Engine and then included in the ADLRBSAS model. The initial filtering and elimination of outliers from these datasets are aided using an application-specific segmentation model. Using 19 Vignette classification models, an improved CNN model is used to evaluate segmented images. The suggested model can classify crop types with 95.4% accuracy, land types with 97.2% accuracy, and water and urban cover types with 95.9% accuracy. The accuracy of the suggested model was evaluated using data from several locations and eras; the results showed that the model was consistently reliable. The accuracy, recall, AUC, and latency of the model were assessed and compared to those of several other state-of-the-art methods. In terms of accuracy, recall, and area under the curve levels, the recommended model outperformed the older ones by 2.4%, 3.5%, and 4.5%.**

Keywords*— *Satellite, Image, Soil, Water, Google-Earth, Engine, Precision, Recall, AUC, Scenarios.

I. INTRODUCTION

Satellite image classification has been a challenge for many different businesses; to do so successfully calls for a concerted effort from all parties, starting with the collection and storage of data and continuing through the extraction and selection of properties. Using models like the one in Figure 1 is a typical approach when classifying satellite photos into several groups. Within these paradigms, a range of pre-processing and classification blocks are made available for various instances

The model first processes some simple satellite photos in the initial step of the process. For application-specific classifications, processing methods including segmentation, fusion, and noise reduction are a few examples. After the pre-processing stage, these images are sent to a black-box feature extraction and classification unit. Convolutional neural networks (CNN) are utilized by this unit for internal training and validations [1, 2, 3].

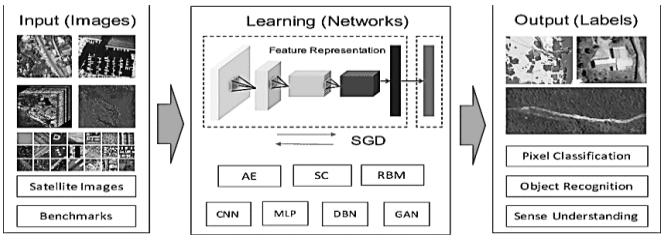

Figure 1. A typical satellite image classification model using deep learning

Once trained, a model can be used to categorize test photos, yielding results that are particular to each circumstance. An overview of these models will be given in the second section of the book, together with an analysis of their key characteristics, drawbacks, and potential future research areas. As has been made evident throughout this discussion, the models under consideration are rather broad, which limits their applicability for application-specific tasks like the classification of satellite images. The third and last section describes how we were able to instantaneously identify satellite photos based on their position using a cutting-edge deep learning model. The work done in this area will raise the value of the earlier sections. We use a variety of metrics, such as classification accuracy, precision, recall, and area under the curve, to evaluate the model's effectiveness which was discussed in Section 4. We assess it by examining all its potential applications (AUC). These parameters are compared with a few alternatives, and state-of-the-art methods to identify any shortcomings in the current model. The final section of this essay presents some insightful critiques of the model that is being used as well as recommendations for improving its overall effectiveness.

II. LITERATURE REVIEW

Scientists have spent a great amount of time and energy developing a broad range of categorization algorithms for satellite photos. These models are often created with a specific goal in mind, with a particular class of satellite images acting as the main focus of their attention. For instance, the study described in [4, 5], and [6] recommends using Multispectral and Multiangle 3-D CNN in addition to Multi-temporal Deep Fusion Networks (MDFN) for the categorization of multi-temporal high-resolution pictures of cities, impervious surfaces, and other objects. These models are constrained in the amount of information they can take in at once since they employ deep learning methods to assess a certain class of picture. To boost the scalability of this capability, the author of "7" advises utilizing a CNN that is equipped with picture models that are sufficiently versatile to be employed in a variety of different remote sensing applications. By expanding the number of applications that may utilize CNN, this capacity can be raised. Models for recognizing land cover, further classifying dominant forest species, and creating a semi-

supervised adversarial deep network are discussed in [8, 9], [10], and [11]. (SSADN). They built their artistic masterpieces based on this strategy. These models aim to make the data less complicated to improve classification accuracy; as a result, they are simple to implement in a variety of real-time application scenarios. The several solutions that various academics have suggested have a lot in common. Using deep learning, these methods include the categorization of land uses and land coverings [11], transformer pretraining for time series classification (TPTSC) [12], artificial training datasets for ship detection [13], and the identification of rice fields [14]. Below is an example of every possible use of the strategy. Before they can be utilized in the process of universal categorization, these notions require further validation since they are used in situations that are particular to a certain area. It follows that scholars would have developed a wide range of classification schemes. However, most of these devices either have a narrow range of applications or struggle in general-purpose settings. In the section that follows, we discuss the steps necessary to construct a successful deep-learning model for identifying application-specific, real-time, region-based satellite images. A neural network will be used to create this model. We were able to show that the suggested model is applicable in a variety of real-world application scenarios by thoroughly evaluating it on various datasets and contrasting it to other machine learning models. As a result, it became clear that the model could be helpful in various situations.

III. Proposed Water and Soil Classification Model

The retrieved satellite images are loaded into an enhanced CNN model to begin the categorization process. This model can handle a lot of data since it uses a range of different CNN approaches. These CNN Models use a variety of techniques, including feature extraction by convolution, variance maximization via max pooling, feature elimination via dropout, and classification using a fully connected network. One of the models being questioned is the VGGNet19, a convolutional neural network model. Figure 2 depicts the culmination of the network model's growth, which is the fusion of numerous CNN layers. The initial resolution of each retrieved image was reduced to 128 by 128 pixels to facilitate the extraction of common features and the development of classifiers. Many convolutional layers are used by feature extraction to complete this operation. These layers' stride, window, and padding sizes may all alter.

In its separate networks, the VGGNet 19 model uses a wide variety of stride values (3x3, 4x4, and 5x5), window widths (8x8, 256x256, and 512x512), and padding sizes (3x3, 4x4, and 5x5). As a result, each satellite image and the areas around it may be evaluated using a tremendous number of qualities, expanding the range of possible representations for the images. The feature activation function found in each convolutional layer, at least in part, regulates the variance of the features. Leaky rectilinear units, or LReLUs, are what is used in this design. Maximizing variance when extracting features is made considerably easier with the help of these modules. The outputs of the several convolutional layers used in the model are controlled by the equation below.

$$Conv_{out_{i,j}} = LReLU\left(\frac{r}{2} + a, \frac{c}{2} + b\right) * \sum_{a=0}^{\frac{r}{2}} \sum_{b=0}^{\frac{c}{2}} I(i-a, j-b) \ldots \quad (1)$$

Where I am the satellite-returned image (or the image obtained from the previous iteration), a and b are the stride sizes (which may vary from 3x3 to 5x5), and m and n are the window widths that the chosen convolution layer uses. During the first filtering and elimination stage, the leaky RELU used in this study deleted 10% of all low-variance features. Equation 2 describes the behavior of the LReLU, as will be shown in the following equation sets.

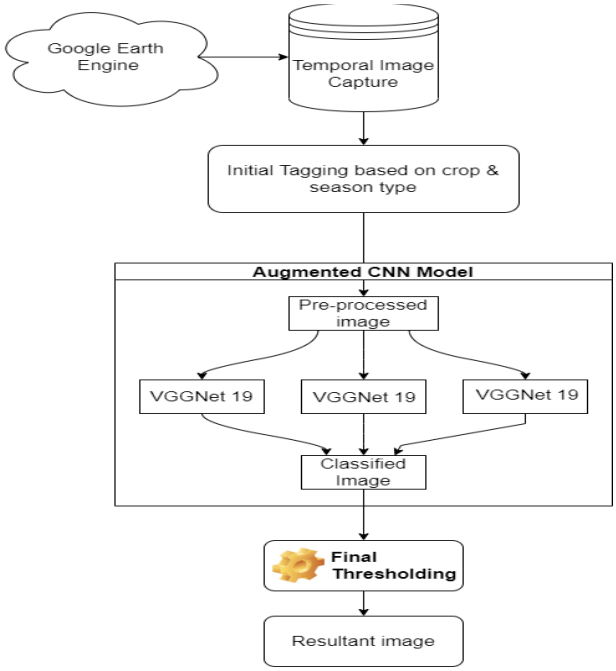

Figure 2. The overall flow of the proposed model for satellite image classification

Where I am the satellite-returned image (or the image obtained from the previous iteration), a and b are the stride sizes (which may vary from 3x3 to 5x5), and m and n are the window widths that the chosen convolution layer uses. During the first filtering and elimination stage, the leaky RELU used in this study deleted 10% of all low-variance features. Equation 2 describes the behavior of the LReLU, as will be shown in the following equation sets.

$$LReLU(x,y) = 0.1 * (x,y), when\ x < 0\ or\ y < 0 else, 1, when\ x \geq 0\ and\ y \geq 0 \ldots (2)$$

The examination of this variance allowed for the large-scale selection and extraction of traits, which created a wealth of new research opportunities. This method iterates with a variety of window and stride widths to predict a broad range of characteristics based on input satellite pictures. Images are classified using a fully connected neural network (FCNN) model following the varied crops and seasons. This is achieved by using the data gleaned from the aforementioned attributes, which is then fed into the model. The images are forwarded to a thresholding layer for further processing once they have initially been tagged. This layer of the image makes it easier to separate a picture into its constituent parts. Every image includes these elements, which show the kind of crops that are growing, the quantity of water that is accessible, the level of urbanization, and the type of terrain that is there. A green component predominates in agricultural regions, a blue component predominates in aquatic areas, a brown component predominates inland areas, and a gray component predominates in urban areas. After the image has been first categorized, each classified pixel is routed via an intelligent thresholding layer to calculate and use class-specific

thresholds for the final classification. This is carried out to ensure that the final categorization is precise. The urban cover is defined as a set of specified criteria that pixels in a picture must not fulfill. Figure 3 provides an excellent illustration of this idea: The components are colored as follows: red for land, green for agricultural areas, and blue for water. The darker hues on the color wheel are intended to be understood as signifying cities.

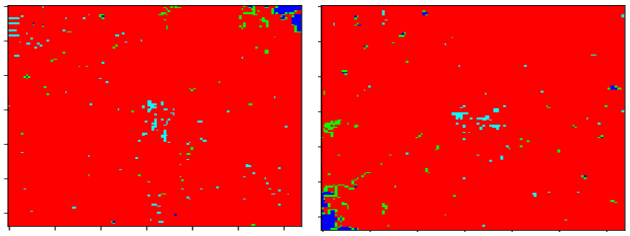

Figure 3. Type of cover for different images (Red represents regions for land, Green represents the crop regions, Blue represents the water regions, and urban areas are represented in other colors)

The cut-offs are established by visually examining different sites over a total of 864 satellite images. These photos were taken across all four seasons and at a wide range of periods of the year. Precision, recall, accuracy, area under the curve (AUC), latency, and other performance metrics were all examined using both the threshold and the classification layer. The area under the curve (AUC) and accuracy were the two other performance measures examined. In the next chapter of this book, the findings of an inquiry into these metrics will be given and discussed in connection with many novel classification techniques for satellite images.

IV. Result Evaluation and Comparative Analysis

The MODIS-006 satellite data sets are available via Google Earth Engine at the start of the proposed model. The data collected are utilized to assess the COPERNICUS-/S1 GRD subset, and the USGS SRTMGL1 was used to determine the ground elevations. The data that was acquired between 2015 and 2021 in a region with a focus on Nagpur was analyzed using the provided model (at 79.1 and 21.1 degrees, respectively, in longitude and latitude). A total of 1,500 images were collected, and they were split evenly across the training, testing, and validation sets (60:30:10). These images, which cover a total of four different climatic zones and two different topography characteristics, were captured throughout each month of the year. Figure 4 displays the findings of a comparison between the accuracy reached with [5] and [9] and the number of test images (NTI) utilized in the evaluation. Both [5] and [9] were used in the assessment. By averaging the accuracy ratings for each distinct crop type, land cover, and water cover region, the total accuracy performance was calculated. This made it possible to evaluate the accuracy more thoroughly.

Given these results, it is evident that the proposed model outperforms [5] and [9] in terms of accuracy across a variety of different domains by at least 20%. We employ a CNN model that has been improved by combining many different high-performance classifiers, which allows us to get better results with lower error rates than we could with more conventional model implementations. We can accomplish this because of this. As can be seen in the image that follows these table sets, the precision (P) values also underwent the same kind of examination.

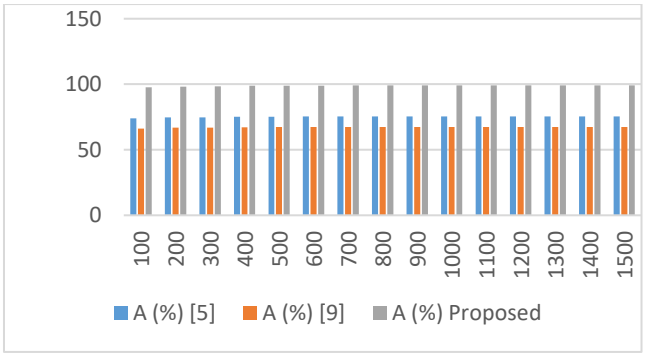

Figure 4. Accuracy of different models for classifications

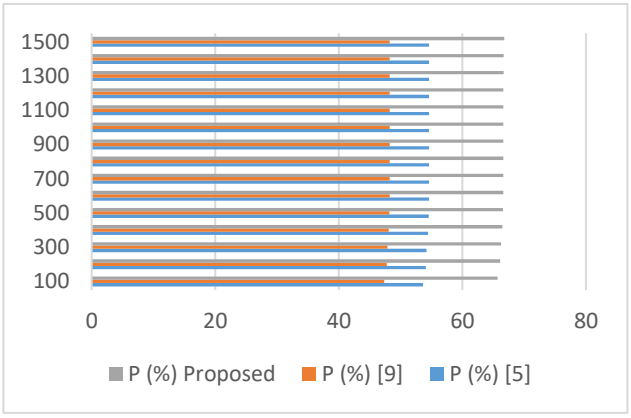

Figure 5. The precision of different models for classifications

The proposed model is 10.5% and 15% more accurate than the R1 and R2 models, respectively, according to satellite images. This suggests that the presented model may be used for a wide variety of real-time applications requiring a high degree of precision. When examining the recall (R) value sets, the results are the same, as shown in figure 6 as follows,

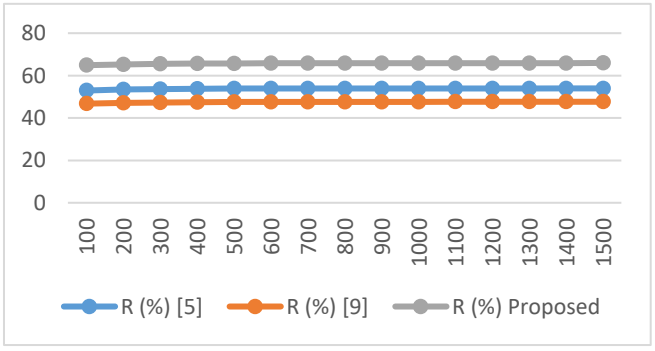

Figure 6. Recall of different models for classifications

It was shown that the recommended model is 12.2 percentage points more successful than [5] and 16.5 percentage points more effective than [9] over a wide range of test satellite images. The created approach may be used in a variety of real-time scenarios that demand exceptional recall due to its versatility. The graphic that follows these gives details that are comparable to the levels of the area under the curve (AUC).

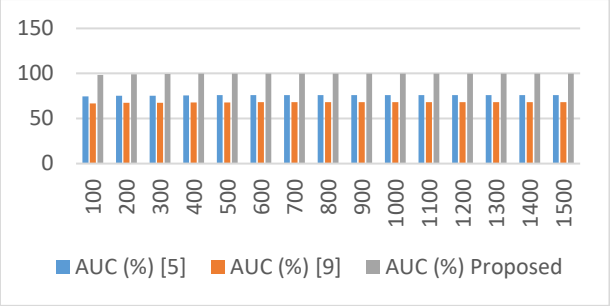

Figure 7. AUC of different models for classifications

The proposed model outperforms [5] by 20.5% and [9] by 26.8% in terms of the area under the curve under a variety of satellite picture conditions. The correlation coefficient of [5] and [9] is. This shows that a range of real-time low-error applications may use the suggested paradigm. This might result in improvements to a variety of satellite image classification use cases' overall classification performance and processing speed for different scenarios.

V. Conclusion and Future Work

The recommended model uses a variety of deep learning techniques, including the VGGNet method, together with a carefully chosen collection of datasets. As a result, the model may be used for a wide variety of tasks that require the classification of satellite pictures. It has been shown that the proposed model outperforms a variety of other models that are currently thought of as state-of-the-art models in terms of accuracy, precision, and recall. The model in issue has an extremely low error rate and almost flawless precision, making it ideally suited for use in real-time applications. The model that is being provided now has a recall rate of 69.5% and an accuracy rate of 66.8% as a result of this improvement. Because of this, it could be helpful in situations where a high level of accuracy and a low error rate are required. The dynamic selection of training and testing datasets enables the proposed model's intelligent feature selection capabilities. This leads to evidence that the proposed model outperforms the vast majority of previously reported classification algorithms, making it suitable for the fast and in-the-moment stratification of satellite data. The researchers believe that by employing Q-Learning, reinforcement learning, and recurrent Neural Networks they will be able to enhance this performance in the future (RNNs). All of these techniques aim to reduce unnecessary effort while achieving modest but discernible improvements in accuracy and effectiveness. To validate the model's efficacy in real-time applications, researchers may also look into new data sources and apply the recommended model to such data samples.

References

[1] H. Yu, H. Zhang, Y. Liu, K. Zheng, Z. Xu, and C. Xiao, "Dual-Channel Convolution Network with Image-Based Global Learning Framework for Hyperspectral Image Classification," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 6005705, Doi: 10.1109/LGRS.2021.3139358.

[2] L. Vidal Batista, "Turbidity classification of the Parafovea River using machine learning and Sentinel-2 images," in IEEE Latin America Transactions, vol. 20, no. 5, pp. 799-805, May 2022, DOI: 10.1109/TLA.2022.9693564.

[3] P. Tang, P. Du, J. Xia, P. Zhang, and W. Zhang, "Channel Attention-Based Temporal Convolutional Network for Satellite Image Time Series Classification," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 8016505, Doi: 10.1109/LGRS.2021.3095505.

[4] A. Constantin, M. Fauves, and S. Girard, "Joint Supervised Classification and Reconstruction of Irregularly Sampled Satellite Image Times Series," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-13, 2022, Art no. 4403913, DOI: 10.1109/TGRS.2021.3076667.

[5] W. Zhang et al., "Attention-Aware Dynamic Self-Aggregation Network for Satellite Image Time Series Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-17, 2022, Art no. 4406517, DOI: 10.1109/TGRS.2021.3120914.

[6] C. Bai, D. Zhao, M. Zhang, and J. Zhang, "Multimodal Information Fusion for Weather Systems and Clouds Identification from Satellite Images," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 7333-7345, 2022, DOI: 10.1109/JSTARS.2022.3202246.

[7] S. Zhao, Z. Zhang, T. Zhang, W. Guo, and Y. Luo, "Transferable SAR Image Classification Crossing Different Satellites Under Open Set Condition," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 4506005, Doi: 10.1109/LGRS.2022.3159179.

[8] C. J. Zhang, X. -J. Wang, L. -M. Ma and X. -Q. Lu, "Tropical Cyclone Intensity Classification and Estimation Using Infrared Satellite Images with Deep Learning," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 2070-2086, 2021, DOI: 10.1109/JSTARS.2021.3050767.