NDR-UNet: Segmentation of Water Bodies in Remote Sensing using Nested Dense Residual U-Net

Jonnala Surekha

nagasurekha.20phd7090@vitap.ac.in

VIT-AP Campus  https://orcid.org/0000-0002-2906-5882

NEHA GUPTA

VIT-AP Campus

Research Article

Keywords: Water-bodies, nested dense residual network, Unet, Digital signature, ECDSA

Posted Date: March 12th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3133244/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
NDR-UNet: Segmentation of Water Bodies in Remote Sensing using Nested Dense Residual U-Net

Naga Surekha Jonnala
School of Electronics Engineering
VIT-AP University, Amaravati, Andhra Pradesh 522237, India
Neha Gupta
School of Electronics Engineering
VIT-AP University, Amaravati, Andhra Pradesh 522237, India
nagasurekha.20phd7090@vitap.ac.in, nehagupta27brs@gmail.com

Abstract
The identification and accurate delineation of water bodies in remote sensing satellite images have significant implications for scientific research and various applications such as natural disaster forecasting, drought and flood detection, and monitoring disappearing water bodies. However, this task poses challenges due to complex spectral variations caused by factors like aquatic vegetation, different colors of lakes/rivers, mud along the sand, and shadows from surrounding plants. To address these challenges and improve water body extraction from high-resolution and moderate high-resolution remote sensing images, we propose a method called D3net (Nested Dense Residual Network). The Adam optimizer is employed to train the satellite images, minimizing the associated losses. The activation function and the number of nodes in each layer are optimized to achieve the best performance. To ensure data integrity and protect the identified water bodies during transmission, we utilize the Elliptic Curve Digital Signature Algorithm as a security component. This algorithm creates a digital signature for the projected area of water bodies, providing data protection. For our study, we used a dataset consisting of 5682 Sentinel-2 satellite images, including 2841 images and their corresponding masks. The masks were generated using the Normalized Water Difference Index (NWDI) for a specific geographic location in Europe. The suggested model achieves a performance IOU (Intersection over Union) of 93.27% and a recall rate of 95.60%. Additionally, the model can be applied to tasks such as edge detection, blurry image recognition, and low-resolution image detection. It exhibits reliability and accuracy in its predictions, although it may require more memory due to the utilization of high-resolution and moderate high-resolution images for segmentation.

Keywords: Water-bodies, nested dense residual network, Unet, Digital signature, ECDSA.

1. Introduction
Accurate segmentation of water bodies is essential for various environmental protection efforts, monitoring of water resources, and assessing the impact of natural disasters. In particular, the detection of changes in water bodies relies on precise and reliable segmentation techniques. This research focuses on addressing the challenges associated with segmenting water bodies in complex and difficult environments using very high-resolution (VHR) and moderate high-resolution remote sensing imagery. Remote sensing images captured by satellites and aircraft provide comprehensive coverage of the Earth’s water surface. However, in VHR and moderate high-resolution remote sensing images, distinguishing the boundaries of water bodies can be challenging. These difficulties arise due to factors such as the presence of aquatic vegetation, the presence of mud or boats along the shoreline, and the shadows cast by tall surrounding plants. To overcome these challenges, advanced segmentation methods are required to accurately delineate water bodies in such imagery. The research aims to develop robust algorithms and techniques that can effectively handle these complexities and provide accurate segmentation results. The proposed methods will contribute to improving the understanding of water bodies, enabling better environmental management, resource monitoring, and disaster response.

In recent years, the utilization of satellite images for the extraction and detection of natural resources, such as water bodies and forests, has had a significant impact. It is crucial to monitor and assess the growth of water bodies to ensure their sustainable development. Monitoring these water bodies serves various purposes, including the early
detection of droughts, floods, and other natural disasters, as well as effective management strategies. Satellite images have emerged as a valuable tool for locating and identifying various elements, including people, animals, buildings, water bodies, and ships. The main objective of this project is to identify and delineate water features using a large dataset of satellite images obtained from Kaggle. To accomplish semantic segmentation, a fully convolutional Nested Dense Residual UNet (NDRUNet), also referred to as D3net, deep learning technique is employed. Various neural network architectures have been explored for image segmentation tasks, but the D3net model demonstrates superior performance with minimal information loss, as highlighted in the study by Ying et al. (2021) [1]. By leveraging the capabilities of the D3net model, this project aims to accurately segment water bodies in satellite images, providing valuable insights for water resource management and disaster response.

In addition to the segmentation of water bodies using the D3net model, the suggested approach incorporates the use of ECDSA (Elliptic Curve Digital Signature Algorithm) to ensure data integrity. The application of ECDSA, a widely recognized cryptographic algorithm, adds an additional layer of security to the detected mask of water bodies. This ensures that the data and identified water bodies remain tamper-proof and maintain their integrity, as highlighted in the innovative work by Ye et al. (2022)

Furthermore, by making minor adjustments to the object detection settings, the suggested model can be utilized to identify Aqua Trash, including debris and pollutants, within water bodies. This adaptation of the strategy enables the model to contribute to the ongoing efforts in monitoring and managing the cleanliness and health of water bodies. By accurately detecting and tracking Aqua Trash, the model aids in maintaining a sustainable environment by mitigating the negative impact of pollutants on aquatic ecosystems. The versatility of the proposed approach allows for its application in various contexts related to water body management and environmental preservation, further emphasizing its potential significance and practicality.

In the pursuit of achieving more accurate semantic segmentation results, particularly in terms of improved edges and boundaries, several researchers have made advancements to the basic prediction outputs. Lin et al. (2017) [2, 3, 4, 5] utilized long-distance residual connections for multi-scale features during the downsampling process to achieve high-resolution (HRL) predictions [2]. Qin et al. (2019) introduced the residual refine module (RRM), an independent encoder-decoder, to post-process the semantic segmentation results and enhance their quality [3]. Yu et al. (2018) proposed the refinement residual block (RRB) to improve feature maps and contribute to more precise predictions [4]. Cheng et al. (2020) developed a dedicated refinery network, CascadePSP, which combines global and local refinement to maximize the accuracy of approximate predictions [5]. It is worth noting that in many of these approaches, the repetitive structural design employed may lead to redundancy or duplication. The main contributions of our work are as follows:

- Proposal of a Simple Multi-Layered Residual Architecture: We introduce a novel architecture called the "nested dense residual UNet" for semantic segmentation. This architecture is designed to enhance the segmentation performance by incorporating nested dense residual blocks. The nested structure allows for efficient feature extraction and preserves important spatial information, resulting in more accurate segmentation results.

- Utilization of the Adam Optimizer: In order to train our proposed model effectively, we employ the Adam optimizer. This optimizer is known for its ability to adaptively adjust the learning rate and momentum, leading to improved convergence and performance during training. By utilizing the Adam optimizer, we achieve better segmentation results and enhance the overall performance of our model.

- Integration of ECDSA for Integrity and Security: To ensure the integrity and security of the predicted mask, we apply the Elliptic Curve Digital Signature Algorithm (ECDSA). This cryptographic algorithm generates a digital signature for the predicted mask, providing a secure means of verifying its authenticity and protecting the data during transmission. The integration of ECDSA adds an extra layer of trust and reliability to our proposed approach.

- Deep Learning Visualization: We employ gradient activation maps to visualize the deep learning process and gain insights into the classification process performed by the network. This visualization technique helps us understand which regions of the input images contribute most to the segmentation decision, aiding in the interpretation and analysis of the model’s behavior.

Overall, our work contributes to the field of semantic segmentation by proposing an effective architecture, optimizing the training process, ensuring integrity and security, and providing insights through visualization techniques. These contributions enhance the accuracy, reliability, and interpretability of our proposed method.

The remaining manuscript is structured as follows: section 2 presents previous state-of-the-art methods; section 3 describes the datasets used in the work; Section 4 gives the proposed method of water body segmentation using NDRUNet; section 5 presents the results of proposed method; and section 6 gives final conclusion of the work.

2. Related Work

Recently, work on DL-based water-body segmentation from remote sensing data has gained attention and
progress. [6][7][8][9][10]. By taking into account both spectral and spatial information, Yu et al. [6] are pioneers in developing a CNN-based technique for water-body extraction from Landsat images. This CNN-based method, however, divided an image into small tiles in order to make pixel-level predictions, which led to a high level of redundancy and low efficiency. A constrained receptive field deconvolution network was suggested by Miao et al. [7] to extract water bodies from HR remote sensing images. To extract water bodies from VHR images, Li et al. [8] used a conventional FCN model. It performed significantly better than techniques using the normalised difference water index (NDWI), support vector machine (SVM), and sparsity model (SM). In order to segment water bodies more precisely, Duan et al. [9] suggested a unique multi-scale refinement network (MSR-Net) that fully used the multi-scale properties. Although the MSR-Net has a multi-scale module, this model does not take into account channel links between feature maps and does not reuse high-level semantic information. Guo et al. [10] introduced a multi-scale feature extractor that included four dilated convolutions with varying rates and was deployed on top of the encoders. They used a straightforward FCN-based technique for water-body extraction. This FCN-based technique did not extract entire features at different scales; instead, it utilised the multi-scale information of high-level semantic features. It is clear that feature extraction and prediction optimization were prioritised in recent CNN-based water extraction experiments, but there is still much potential for improvement.

In order to identify minor water bodies and conduct quality analyses in Wuhan, Teng Fei et al. [11] developed a model. The approach is based on band analysis and image categorization. For quality analysis, a semi-empirical deep learning algorithm is employed. Since it is entirely based on an inversion model, accuracy evaluation is performed at regular intervals then finished by the inversion of water quality. The methodology has certain flaws, including the difficulty of detecting water features in mixed images.

A technique for segmenting water bodies from remote sensing images using mask R-CNN is suggested by Fengyu Yang et al. [12]. The proposed method addresses the drawbacks of inadequate resolution and pixel clarity in remote sensing images. The model was created using Google Earth images as a foundation. The dataset is adjusted by data augmentation. Both the ResNet-50 and the ResNet-101 techniques were used to train the model. The two main strengths are (i) the predictions’ excellent accuracy and (ii) the model’s high viability and adaptability.

To solve a number of issues in the field of remote sensing, Jorge Lira et al. [13] offer a technique. This model makes use of optical reflectances and segmentation methods. The water bodies are separated out of the satellite images using a unique technique called PCA. The remote sensing satellite image is split into two halves using the clustering technique, depending on pixels. The water body is shown in the first part, followed by the remainder of the satellite image. The above approach has a problem in that false overlays is shown at various points during detection.

R. Lalchhanhima et al. [14] proposed a method. That use a hybrid CNN that takes into account U-Net and Inception, makes use of sparsely distributed SAR data, and speckle noise image segmentation. The output of this model depends on a number of aspects, including the training, dataset size, and accuracy, among others. As the kernel size increases, training time increases. To create a binary image of 128X128 pixels, Otsu’s threshold approach is used to the resultant image. The model demonstrates its superiority by (i) outperforming existing CNN techniques and, (ii) the same level of performance can be attained with smaller train data and less training time.

Due to a variety of variables, Zhili Zhang et al. [15] discussed the difficulties in precisely determining the limits of water bodies. A new combination and multi-feature extraction are provided as solutions to these problems. To get rich feature representation, three feature extraction sub-modules are employed for different spaces and connections. The multi-scale detection fusion module is employed simultaneously for contour detection. Encoding-decoding semantic feature fusion module is further utilised. It has two benefits: It performs cutting-edge segmentation; and (ii) the model is sturdy.

A change for the current ECDSA and their uses was made by Yuanbo Shnag et al. [16]. The ECDSA’s efficiency and security are enhanced by this work. The ECDSA method is chosen because it has fewer drawbacks and more advantages than the RSA one. It increases effectiveness by analyzing various assaults, such as side-channel attacks. By increasing its effectiveness and defending it from numerous attacks, this activity has its own benefits.

A technique, verifiable ECDSA, was introduced by Xiao Yang et al. [17]. The main objective is to be able to use a public key provided by the reliable service to encrypt the digital signature. This development has added new capabilities to the current VES scheme. It may have numerous signatures with the same structures or those that are similar but have different variations. Later, it was adopted by many crypto currencies. The biggest disadvantage of this proposed work is the storage need of a 270 MB assessment look-up table.

A model based on Mobile-UNet was created by Jing et al. [18]. To find fabrics, it use the CNN approach. Segmenting fabric flaws from end to end can be done effectively. Simple imbalance is reduced by applying the loss function of the median frequency. The encoder component is built under MobileNetv2, and five layers of decoders are added. Using a softmax layer, segmentation mask detection is performed. The model’s cutting-edge precision is one of its main advantages.

Tensorflow was used by Wang et al. [19] to give the scientific computation of fluid flows. In this project, the graph-based module of Tensorflow is utilized. Additional performance and accuracy analyses revealed great scalability up to TPU v3 pod. The TPU platform is used to
create the model. A grid search for solar power forecasting with LSTM using Nadam Optimizer was found by Jatin Sharma et al. [20]. The LSTM provides better results for time-series data. In this study, eight alternative neural network models and an optimizer are compared against the LSTM along with two time series. Each optimizer’s accuracy is determined, and the results are compared. Nadam with LSTM ultimately demonstrates more accuracy.

A research of semantic segmentation for aquatic bodies was introduced by Erfani et al. [24]. The ATLANTIS dataset is the foundation of this effort. This might stop emergencies from occurring during floods. The dataset includes several classifications across numerous groupings. For the purpose of distinguishing between aquatic and non-aquatic environments, the AQUANET model was created. The strongest part of the work is its dataset, which can cover a wide range of classes with different types of dataset. Using high quality satellite images, Rajyalakshmi et al. [22] established a method to identify water bodies. Thresholding methods are used in this paper. The dataset consists of images with a wide range of spectral and temporal characteristics. In this study, a brand-new technique called single-band threshold employing bilateral filtering is used. The primary restriction on this study is that only HRL images may be utilized.

A technique for IOV with a fault-tolerant ECDSA signature was proposed by Lin et al. [23]. Moreover, it is vulnerable to various attacks. Pedestrians and vehicles can communicate with each other. There is a possibility that a third party can view this sensitive information during transmission. A fast, fault-tolerant ECDSA protocol is proposed to defend against the following attacks.

A U-net has been proposed by Fang Chen et al. [24] as a building identification technique. Using remote sensing photos, this model helps identify buildings. It is created by replacing the n-channel feature maps with a series of smaller meshes. In addition, an edge loss function is implemented to improve the recognition performance.

On a multi-spectral image with 13 bands, Gordana Kaplan et al. [25] describe a process that combines the usage of a pixel-based index and an object-based strategy. Sentinel-2 images with resolutions ranging from 10 meters to 60 meters are used in the model. The model considers a mountainous and urban setting for a better understanding of the performance. Using NWDI, the indexes are retrieved. A straightforward pixel-based approach is used to incorporate additional methods, producing a low albedo as a consequence. The model’s additional benefits include (i) accurate water monitoring in a variety of climates and geographies and (ii) enhanced water extraction due to the integration of pixel- and object-based categorization. The model’s disadvantage is that it utilises only two of the 13 available bands.

In all above mentioned methods used limited datasets and achieved less accuracy. The mentioned models did not have the security for predicted masks and computational cost is also more. So, we have developed NDRUNet model for the segmentation to achieve the more accuracy and ECDSA is used to protect the predicted mask details.

3. Materials

In our work, we collected Sentinel-2 aerial images of water bodies in the Europe region. Each image is accompanied by a corresponding black-and-white mask, where black represents non-water areas and white represents water bodies. The masks were generated using the Normalized Water Difference Index (NDWI), a commonly used index for vegetation identification in satellite images. However, we used a higher threshold to specifically identify water bodies.

The training data for our DL model was obtained from publicly available datasets [26], which provided 2841 masks corresponding to the 2841 Sentinel-2 satellite images [20]. These satellite images were captured by the Sentinel-2 satellite and have spatial resolutions ranging from 10 meters to 60 meters [27, 28]. The dataset consists of two directories: one for the images and another for the masks. The masks are binary, with white indicating non-water areas and black indicating water bodies.

It’s worth noting that the Normalized Difference Water Index (NDWI) [29] can refer to different water-related indexes derived from remote sensing. One such index, proposed by Gao in 1996, utilizes near-infrared (NIR) and short-wave infrared (SWIR) wavelengths to monitor variations in leaf water content [30]. While NDWI is mentioned here, we specifically used the Normalized Water Difference Index (NWDDI) for generating the masks in our work. This can be represented mathematically as

\[
NDWI = \frac{W_{NIR} - W_{SWIR}}{W_{NIR} - W_{SWIR}}
\]

where, \(W_{NIR}\) and \(W_{SWIR}\) is the reflectances of the near infrared band and reflectances of near short-wave infrared reflectances respectively.

\[
NDWI = \frac{W_{G} - W_{NIR}}{W_{G} - W_{NIR}}
\]

where, \(W_{G}\) is reflectances near green bands. Equations [1] and [2] both are used to monitor the changes in the water level. This NDWI creates the digital and visual interpolation of the output image that is from -1 to 0 shows bright surface and no evidence of the water content and +1 represents the water content.

4. Method

The data used in our proposed methodology is sourced from the Water Bodies dataset provided by [26]. This dataset comprises of masks that are applied to the corresponding original images. The dataset does not include any additional parameters, only the images themselves.
In developing our model, we followed an incremental approach based on the agile development methodology. This approach allows for flexibility in incorporating future enhancements and maintenance. The model construction process is divided into seven key high-level phases, which help guide the development and ensure a systematic and efficient approach to building the model.

During the initial stages of our methodology, the dataset is fed into the model, and exploratory data analysis is conducted to gain insights and identify any irregularities that may impact the model’s performance. This step is crucial to ensure the smooth functioning of the model. As part of the exploratory data analysis, a comparative visualization of the original images and masks is performed. This helps in identifying any discrepancies or differences between them. In some cases, inconsistencies or contradictions may be observed in the ground truths provided by the masks. To address this, further investigation is conducted to check for any rotations or padding in the images that could have caused the discrepancies. In the data preprocessing step, the images undergo transformations to prepare them for model training. One of the initial transformations applied is resizing. Some images may have larger dimensions, making them difficult to handle in memory and significantly different from other images in the dataset. To address this, a threshold of 1500 is set for both the height and width of the images. Any images exceeding this threshold are resized or re-scaled to meet the specified dimensions. By resizing the images, they are made more manageable and brought to a standardized size, facilitating further processing and analysis in subsequent stages of the methodology.

After the resizing of the images, the next transformation operation is the mask conversion. The original images are analyzed by displaying their histograms, and it is observed that the majority of the pixel values are centered around zero. This indicates the possibility of rotation in some of the images. To address this, the brightness of the values in the middle range is increased using histogram equalization. This helps to enhance the visibility of the image details. In the case of masks, they need to be converted into a binary format with distinct values representing the presence or absence of water. This is achieved by applying a thresholding technique, where a specific threshold value is used to separate the pixels into two distinct classes. The resulting binary masks are then saved in the JPEG format. The third transformation function applied is scaling. The main objective of this transformation is to accelerate the convergence of the model during training. To achieve this, the original images are normalized or scaled to the range of [0, 1]. However, it is important to note that the mask images are not taken into consideration during the scaling process, as they have already been converted into binary format and do not require scaling. By performing these transformation operations, the data is preprocessed and prepared for further processing and training of the model. The resizing, mask conversion, and scaling opera-
tions help to standardize the data, enhance the visibility of image details, and improve the convergence of the model during training.

Padding is an important transformation operation that is applied to the images as the last step of data preprocessing. In the context of neural networks, padding refers to adding extra pixels or values around the borders of an image. The purpose of padding is to ensure that the size of the input and output images remains consistent throughout the network layers, and to enable the down-sampling process to properly divide the image size based on the specified strides. When working with neural networks, the network architecture often requires a specific connectivity pattern and expects the input and output images to have the same size. By adding padding to the images, more connections can be established between non-consecutive layers, and the size consistency is maintained. This helps to ensure that the network can effectively capture and process information from different parts of the image. The padding operation allows for a consistent representation of images, irrespective of their original size, and enables the neural network to effectively learn and extract features from the input data. It ensures that the model can handle images of any size and maintain the desired connectivity patterns, ultimately contributing to the overall performance and accuracy of the model.

In the training phase, the proposed model utilizes the NDRUNet architecture for segmentation. NDRUNet has been shown to outperform other techniques in the task of segmentation. The model is initialized with normal distribution for the weights, ReLU activation for the hidden layers, and sigmoid activation for the output layer. To construct the model using NDRUNet, convolution, encoder, and decoder blocks are built. The convolution block is a key component and has two important considerations. The first consideration is the filter count, which is set to 64 by default, and serves as the input for both the convolution layer and another input. The suggested model employs partial convolution, which can be beneficial for handling missing or occluded data. The encoder and decoder blocks are constructed straightforwardly. The encoder takes inputs from top to bottom and consists of two convolutional layers followed by ReLU activation and batch normalization. The decoder block requires inputs, a skip connection, and a filter count. The input is up-sampled to match the skip connection size, and the previous input is concatenated with the up-sampled input to obtain the final skip connection value. By building the model using the NDRUNet architecture and considering the specific configurations of convolution, encoder, and decoder blocks, the model is designed to effectively learn and capture the features necessary for accurate segmentation of water bodies in the remote sensing images.

The methodological structure for the model detection of water bodies utilizing NDRUNet and tensor flow with ECDSA is shown in Figure 3. It gathers the dataset, goes through a number of intermediate phases, and then produces the original image together with the anticipated mask and an ECDSA digital signature. The steps involved to process the proposed model are:

- Collection of Dataset.
- Splitting the Dataset into train and test groups with a 70:30 ratio.
- Data Processing.
- Building NDRUNet.
- Train the model by defining training parameters.
- Testing the model and calculating IOU.

4.1. Data Processing

In this case, the data processing step focuses on cleaning up the data and removing any irrelevant or unnecessary information. It ensures the dataset is consistent and ready to be used as input for the model. This may involve removing trash or unused information from the data. Additionally, normalizing the data is an important part of the preprocessing step. This involves reducing the fill bits and scaling the pixel values of the images to a specific range, such as [0, 1]. Normalization helps in achieving better results during the training and testing phases of the model. The objective of this data preprocessing step is to prepare the dataset in a standardized and suitable format for the subsequent training and testing steps. It ensures that all the images have the same size and are properly processed, eliminating any irrelevant components that may affect the model’s performance.

4.2. Building blocks of NDRUNet

The visual representation of the UNet architecture is shown in (Figure 4) and the building blocks of the NDRNet is shown in (Figure 5). The proposed architecture of NDRUNet is a combination of both UNet and NDRNet. The details of the NDRUNet model are shown in Figure 6. NDRNet incorporates multi-dilated convolutions using the D2-block within the network model. The total number of parameters in the NDRUNet architecture...
is 2,006,337, and all of these parameters are trainable. It appears that NDRUNet combines the strengths of both UNet and NDRNet to achieve better segmentation results for water bodies. The specific architectural details and parameter counts provide insights into the complexity and capacity of the model.

4.3. Adam Optimizer

The adaptive moment estimation (Adam) optimizer is an algorithm used for gradient descent optimization. It is particularly effective when dealing with large problems that involve a substantial amount of data or parameters. The Adam optimizer combines the Root Mean Square Propagation (RMSP) algorithm and the Gradient Descent with Momentum method. In the proposed model, the Adam optimizer is utilized with specific settings. The model is trained for 30 epochs, with a learning rate of 1e-4. Additionally, a learning rate drop factor of 0.3 is applied. These settings are chosen to optimize the model’s performance during the training process. The Adam optimizer is known for its ability to handle large-scale problems efficiently while requiring relatively low memory usage. It has been widely adopted in various deep learning applications due to its robustness and effectiveness in optimizing neural networks. Comparing different optimizers, the Adam optimizer has been found to provide better performance in the proposed model for water body segmentation. 

4.4. ECDSA

In order to enhance the integrity of the predicted mask in the proposed deep learning model, the Elliptic Curve Digital Signature Algorithm (ECDSA) is applied. ECDSA is a cryptographic algorithm that offers advantages in terms of power and scalability compared to other algorithms like RSA. ECDSA utilizes keys derived from elliptic curve cryptography, which allows for shorter key lengths while achieving the same level of security as other digital signature algorithms. The final digital signature is generated by converting the signature key into bytes. For example, the NIST192p curve, which has a length of 24 bytes, can be used to obtain the signature key. To verify the digital signature on the receiving side, the verification key can be derived from the signing key. This ensures the integrity of the predicted mask and provides a mechanism for validating the authenticity of the generated results. By employing ECDSA in the proposed model, the integrity and security of the predicted mask are reinforced, enhancing the overall reliability of the deep learning-based water body segmentation system.

5. Results and discussion

We have implemented this model using 32GB RAM, 1TB SSD with NVIDIA Graphics desktop in Google co-lab. We have used 5,682 images consists of 2841 satellite images and their respective masks. These images were fed into the proposed NDRUNet model to do semantic segmentation. The training and testing curve of the model is shown in the figure 7. Figure 8 shows the output of the proposed model and shows the original images, respective actual mask of images and predicted images. In this figure, water source is clearly visible and differentiated with other objects present in the images. The performance metric of the model is tabulated in the table 1. After classifying we have calculated the Intersection Over Union (IOU) using the mathematical relation

\[
IOU = \frac{T_p}{T_p + F_p + F_N}
\]

where, \(T_p\) is true positive, \(F_p\) is false positive, \(T_N\) is true negative, and \(F - N\) is false negative. We have achieved an good accuracy of 93% of IOU.

To understand the network behaviour for classification we have applied the Gradient Activation Map (GCAM) to images and masks. This is shown in the figure 9. The darkened color portion are shown that portion is more needful for detection and more features are used for the classification. Lightened color portion of image shown that portion is not more useful for classification and minimum number of features are used for classification.

Finally to validate our proposed model, we compared our model to state-of-the-art models. Table 2 shows the comparison of our model with other models. In [36] they used basic UNet architecture to segment the satellite images with higher accuracy of 85.58%. In [2] used RefineNet with Resnet101 pretrained convolutional neural network (CNN) for classification and achieved the IOU of 93.27%.
**Figure 5:** Nested Dense Residual Network (NDRNet) Architecture

**Table 2:** Comparison of Proposed Method with Previous State-of-the-art models

<table>
<thead>
<tr>
<th>Model</th>
<th>Method</th>
<th>Backbone</th>
<th>IOU(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Net</td>
<td>—</td>
<td>—</td>
<td>85.58</td>
</tr>
<tr>
<td>RefineNet</td>
<td>Resnet-101</td>
<td>86.21</td>
<td></td>
</tr>
<tr>
<td>DeeplabV3+</td>
<td>Resnet-101</td>
<td>86.50</td>
<td></td>
</tr>
<tr>
<td>DANet</td>
<td>Resnet-101</td>
<td>87.90</td>
<td></td>
</tr>
<tr>
<td>CascadePSP</td>
<td>DeepLabV3+&amp;Resnet-101</td>
<td>87.00</td>
<td></td>
</tr>
<tr>
<td>MECNet</td>
<td>MEC + MPF + DSFF</td>
<td>90.64</td>
<td></td>
</tr>
<tr>
<td><strong>Proposed</strong></td>
<td>NDRUNet</td>
<td>D3Net+Unet</td>
<td>93.27</td>
</tr>
</tbody>
</table>

86.21%. DeeplabV3 with Resnet101 is used for segmentation and prediction in [57] and they achieved an 86.50% of mean IOU. In DANet [38] with Resnet101 CNN model achieved the mean IOU of 87.90%. CascadePSP has achieved an accuracy of 87.00% that is combination of DEEPLABV3 and Resnet101 CNN. [15] used MECNet is combination of MEC + MPF + DSFF for classification and achieved 90.64% of mean IOU. Our proposed work used the combination of UNet and NDRNet for the segmentation and classification achieved highest IOU of 93.27%, 95.60% of accuracy, 90.88% of precision, and 86.90% of recall.
6. Conclusion

This study presents a deep learning model based on NDRUNet and TensorFlow for water feature identification in satellite images. The model demonstrates strong performance even on small and hazy satellite images. The architecture of NDRUNet specifically caters to water bodies near land boundaries, ensuring accurate segmentation results in such scenarios. The model utilizes the Adam optimizer, which contributes to faster computation times compared to other optimization algorithms. Data preprocessing steps, such as scaling, mask conversion, and padding, are carefully followed to ensure data consistency and improve model performance. The proposed NDRUNet architecture leverages multiple channels and feature maps, along with appropriate kernel initializers, output activation functions, and hidden activation functions. Max-pooling is employed
to address image dimensionality and enhance computational efficiency. The model excels in accurately identifying water areas, even in challenging weather conditions. To enhance security, the identified water areas are protected using ECDSA, and a secure signature is generated to prevent unauthorized access. Future extensions of this model include the incorporation of video input and utilizing cloud storage for storing observed water locations. The model achieves impressive performance metrics, including an IOU (Intersection over Union) of 93.27%, accuracy of 95.60%, precision of 90.88%, and recall of 86.90%. Overall, this deep learning model presents a robust approach for water feature identification in satellite images, offering improved accuracy, efficiency, and security.

7. FUNDING

The authors state that this work has not received any funding.

8. Data availability

The dataset used in this work is available on Kaggle as Satellite images of water bodies: https://www.kaggle.com/datasets/franciscoescobar/satellite-images-of-water-bodies

Conflict of Interest

There is no conflict of interest.


