Kuwait Journal of Engineering Research
ISSN: 2984-8016
Vol 1 Issue 1 (2023)

Design and Analysis of Improved Deep Convolutional Neural Network Based Satellite Remote Sensing Image Segmentation System

Manisha Goyal[1*], Dr. Nikita Jain[2]

[1]Research Scholar, Department of Computer Science, Poornima College of Engineering, Jaipur

[2]Associate Professor, Department of Computer Science, Poornima College of Engineering, Jaipur

[*]Emails: git.manisha@gmail.com

Submitted: 20/01/2023 **Revised**: 24/02/2023 **Accepted**: 15/03/2023

Abstract: In recent years, machine learning techniques have shown promising results in image segmentation tasks. Machine learning algorithms can learn complex relationships between input features and output labels, and thereby, improve the accuracy and efficiency of segmentation. There are several machine learning algorithms that can be used for image segmentation, such as supervised, unsupervised, and semi-supervised learning. Supervised learning involves training a model on labeled data, where each pixel in the image is assigned, a label indicating the class it belongs to. Unsupervised learning, on the other hand, does not require labeled data and groups pixels into clusters based on their similarity. Semi-supervised learning is a combination of both supervised and unsupervised learning, where a small portion of labeled data is used to guide the clustering process. Or also label them. In the first case, some implementations do not require training. This research aims at developing improved satellite image segmentation technique for segmentation and labeling of satellite images based on machine learning and deep learning techniques. The results have been compared on the basis of figure of merits such as precision, recall, f1 score and analysis proves that the proposed hybrid technique outperforms traditional machine learning and deep learning techniques. The research may be used for analysis and applications related to surveillance, remote sensing and analysis of important parameters such as deforestation and climate change.

Keywords: Supervised Learning, Unsupervised Learning, Segmentation, Satellite Image, Precision, Recall, F1 Score.

1. INTRODUCTION

Over the last decades, a lot of research has been done in the field of computer vision. Image processing algorithms have been developed, aiming to “translate” the actual content of an image, from a matrix of binary values, into meaningful information, from a human point of view. The computer cannot understand and interpret the content of an image by its own. Besides generic image transformation algorithms - like noise removal, color filtering, and pattern recognition – there is a more challenging direction of study: image segmentation and its applications in a semantic context – translating the picture into semantically meaningful, classifiable content. Such algorithms are useful for search engines, robots (automated driving, interaction etc.), medical research and others.

Color image segmentation is defined as a process of extracting from the image domain one or more connected regions satisfying uniformity (homogeneity) criterion which is based on feature(s) derived from spectral components.[8] Segmentation algorithms take an image as input and compute a set of regions, built according to specific similarity/dissimilarity criteria. The resulted regions must form an optimal split of the whole image. What exactly is meant by “optimal” depends on multiple criteria and is still an arguable fact. Depending on the application and the area of study, segmentation approaches can take into account similarity criteria based on color, texture, shape, brightness (or, eventually, a combination of these).

The ultimate goal of segmentation is, in any case, to transform numeric or binary data into meaningful content from the human perceptual point of view. Here we get to the notion of *image understanding*. One popular use of such technology would be image search. At this time, search engines are unable to actually use image information to retrieve best results. Image searches are actually textual searches: the engine looks for websites containing the search term and fetches images embedded in those pages. Other method is search by image tags, which is the same thing – manually tagging images requires a lot of time and labor, is prone to errors and exhaustive on large image databases.

There are multiple technically approaches to segmentations. In [8], four main types of segmentation are defined. They refer to the actual processing of the image file in terms of relevant units:

- Pixel based segmentation – pixels are taken into account separately, one by one, and features such as color, brightness, intensity are considered at pixel-level. Based on this analysis, the pixels are grouped using clustering algorithms, such as K-means, Nearest Neighbor, Hierarchical trees and others.
- Area based segmentation – usually working on groups of pixels (also called *superpixels*). Such methods reduce computational time for large images, because instead of checking millions of pixels, the algorithm checks only hundreds or thousands. Same features are considered, excepting this time sets of pixels are taken as a whole and extended/trimmed according to average/median values of various features for all pixels in the region. It can also use clustering algorithms.
- Edge based segmentation – an edge is defined by a high dissimilarity between two sets of pixels over one ore many criteria. It can split objects that have high intensity/texture/color variation or light glare.
- Physics based segmentation (or semantic segmentation) – it is a sort of combination of all the above, with the final purpose to detect actual objects, despite any color or intensity inner variation. It's the trickiest of all, requires a lot of data analysis (for object shape/position matching and recognition) and machine learning algorithms. It's computationally expensive, but the great advantage is its capacity to learn, thereby including artificial intelligence.

The purpose of this work is to explain how semantic algorithms work, review some implementations with their advantages and drawbacks, to make comparisons based on results and execution time, and highlight optimal approaches, according to the problem. There is also an important conceptual part to be discussed for a better understanding of how the human visual system works and how the brain processes the plain visual information and transforms it into meaningful content. For this, some psychology principles need to be discussed, because apparently it's not enough to resume to mathematical approaches. The final goal is to obtain an intelligent system capable to reproduce the human understanding of visual scenes and for that we need to start from the same premises and provide the input in a more meaningful, human-like way, not only in numbers. Image classification is a vital component in remote sensing, image scrutiny and pattern identification. The image categorization thus figures out a significant tool for inspection of the digital images. Digital image processing involves the process of digital images by using digital computers. Image preprocessing is one of the techniques for the purpose of enhancing the satellite images before the formation of computational process. Image filtering refers to one of the preprocessing technique which applies varies effects on the satellite image to reduce noise and enhance the quality of an image.

2. OBJECTIVES AND SCOPE

1. To design and develop machine learning based segmentation of remote sensing images.
2. To develop improved framework for design of classifier and feature extraction methodologies.
3. To design and develop the application of remote sensing in monitoring deforestation and climate change analysis.
4. To compare the accuracy and performance of segmentation based on performance parameters like precision, recall and f score.
5. To compare the performance of proposed algorithm with contemporary algorithms from the available literature.

3. LITERATURE REVIEW

Sultana et al. (2018) The authors have assumed similar subtleties in composition of some of the best model groups, examining the superiority of their preparation so as to provide better understanding of the tuning of these models by hyperparameters. Chen et al. (2018)]In this paper, it is possible to explore terrible convolution, a powerful way of specifically altering the chain's field of view and to monitor feature resolution in the use of semantine Image segmentation calculated by Deep Convolutionary Neural Networks. We build modules with the usage of atrocious turbulence in a waterfall or to record multi-scale scenarios simultaneously, taking several atrocious rates to segment artefacts in several scales. We propose to extend the Atrous Space Pyramid Pooling Module, which already proposed, which samples multi-level convolutionary features encoding global contexts and improving output. Our training experience also focuses on the integration and sharing of information. The suggested DeepLabv3 framework significantly enhances our previous DeepLab release without Dense CRF post-processing[105] compared to other sophisticated benchmarks for the PASCAL VOC Segmenting Benchmarking of PASCAL VOC 2012. Zhao et al. (2018) This paper proposed an improved semantic image segmentation process based on super pixels and conditioned random fields. The method proposed will take full advantage of the super pixel edge knowledge and the constraints of various pixels. First, we use completely innovative networks (FCNs) to obtain semantic pixel-level functionality and use simple linear iterative clusters (SLIC) respectively to generate super pixel-level region information. Then by fusion of the result pixel level and super pixel level the segmentation results of the image borders are optimized. Finally, to further enhance semantic segmentation accurately, we use the color and location information of pixels using the pixe-level prediction capability of CRFs. In short, this better approach has advantages both about excellent extraction and reasonable adherence to boundaries. Experimental findings on both PASCAL VOC 2012 and Cityscapes datasets demonstrate that in comparison to the conventional FCN model, the proposed method could lead to significant improvements in segmentation accuracy. Zhang et al. (2018) The Deep Image-to- Image arrangement (DI2I) is first developed on X-beam like Digitally Reconstructed Radiographs (DRRs) from 3D CT volumes, particularly for multi-organ divisions. Propose a model system for learning programmed X-beam Image parsement from named 3D CT images. The authors create a Task-Oriented Generative Adversarial Network (TD-GAN) to synchronize and parse unnoticeable genuine X-beam Images at this point. No comment from the X-beam Image area is needed in the entire model pipeline.

Nakazawa et al. (2018) Using deep convolutional encoders to describe and fragment abnormal wafer deformities – neural decoders. The authors construct wafer models for 8 concepts of concept imperfection that are used to plan, validate, and verify data sets using a model of the deformation design period. One of the key capabilities for any oddity detection device is to distinguish unsightly instances. The author's show that models can be retrieved using only the built wafer maps which have the planning concepts. This ability to identify irregular signs without using genuine wafers to prepare a collection of data is helpful because travel occasions occur occasionally. The approach proposed can be used as the basis for separating the example and for subsequent analysis of data, for example, by using standard image preparation techniques and AI methods to get important information, during the process of innovation growth, designers must recognize weak community qualities other than knowing the bunch groups in deformity. Liu et al. (2019) Semantic image segmentation has been used in a variety of areas, including medicine and intelligent transportation, and has developed into a critical application in the field of image processing and computer vision. Researchers publish many data sets in order to verify their methods. For several years, semantic segmentation has been studied. Segmentation has made considerable progress since the development of the Deep Neural Network (DNN). We divide semantic image segmentation methods in two groups in this paper: traditional and new DNN process. First, we summaries briefly the conventional system and data sets that have been released for segments and then analyses extensively recent DNN methods defined in the eights: completely convolutionary network, sample way, FCN in combination with CRF methods, dilated revolution approaches, backbone network advances, pyramid methods. In conclusion, a conclusion is drawn in this area. Chatterjee et al. (2019) They extended to three sectional semantic divisions from the distance sensor awareness. The authors have used a new approach to combine the best quality of classes in entirely convolutionary networks such as Dense squares and to re-calibrate the use of SE squares. The writers have discerned and combined the choices concerning genuine qualities and opinions for the specific undertaking. The authors showed that the engineering proposed beats other methods for class counting. In various presentations of planning, the proposed ICT network was used for drawing parallel maps of various command correctnesses which were then used to eliminate limitations. Hatamizadeh et al. (2019) Suggest limited CNNs for clinical images division. The networks are built to display organ-limited data, both through a specific branch edge and edge-conscious conditions of misfortune and a trained start to completion. The authors conclude that the BraTS 2018 dataset is sufficient to split the mind tumour errand. The work shows that the method results in more and more trustworthy divisions, making a progressively broad application to clinical segmentation promising. The authors decided to use the BraTS 2018 data to make the approach to the division of mind tumours feasible. The conclusions indicate that, in comparison to well-known U-Net and V-Net networks, the system produces increasingly precise divided yields with ne-grained limits. Liu et al. (2019) A proposal in view of the cell-level structure framed by a radical design search zone to look at the system-level structure; The authors present a research area on a system level, which combine different frameworks and generates a concept that allows expert engineering research (3 P100 GPU days on cityscapes). They show that the planned Cityscapes plan, the PASCAL VOC 2012 datasets and the ADE20 K datasets can be rendered feasible. Auto-DeepLab, engineering specially searched for the semantic image division and achieved state-of-the-art effectiveness without pre-training ImageNet. Stan et al. (2019) The method of measuring the thick induction for CRF is converted into CNN operations, which are then defined as layers of repetitive neural networks. Shows that presentation can be enhanced by preparing NNs with a greater number of littler images checked with the fixed information preparation calculation. The optimal image size and sum of the image preparation are recognized for the XCT and SS data sets. Furthermore, the application of the most notable XCT and SS-NNs in related datasets has evaluated NN transmission. While the underlying divisions have been successful, fundamental changes to the raw images have improved NN execution.

4. METHODOLOGY

There are several popular machine learning algorithms used for image segmentation, including k-means, support vector machines (SVMs), random forests, and convolutional neural networks (CNNs). K-means clustering is a commonly used unsupervised learning algorithm that partitions pixels into k clusters based on their similarity.

Figure 4.1: Semantic Segmentation v/s Object Detection

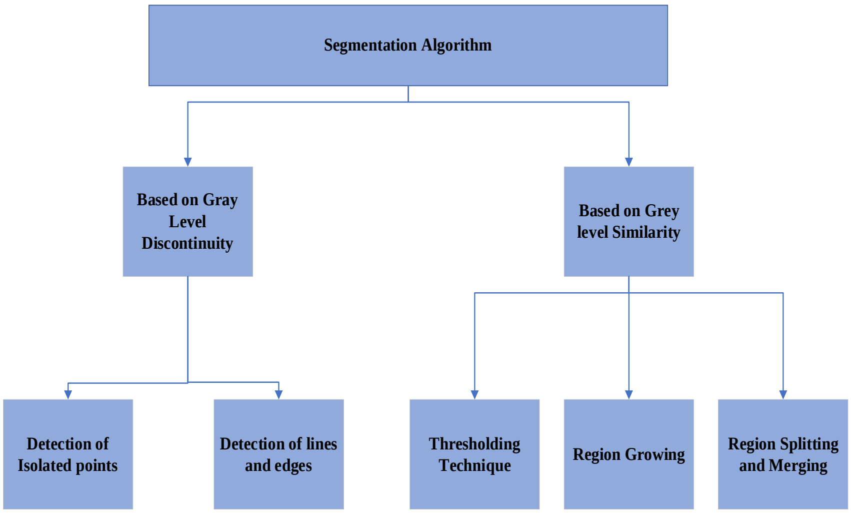

Figure 4.2: Classification of Segmentation Algorithm

SVMs are a supervised learning algorithm that learns a decision boundary between different classes. Random forests are a tree-based ensemble learning algorithm that combines multiple decision trees to improve the accuracy of segmentation. CNNs are deep learning algorithms that learn hierarchical representations of input images and are known for their superior performance in image segmentation tasks.One of the main advantages of using machine learning for image segmentation is its ability to handle large datasets efficiently. Remote sensing images can have millions of pixels, and manually labeling them can be time-consuming and error-prone. Machine learning algorithms can learn from large datasets and make accurate predictions on new data. Another advantage is the ability to incorporate additional features such as texture, color, and shape, which can improve the accuracy of segmentation.

4.1 Methodologies of Remote Sensing Image Segmentation

Image segmentation is the process of dividing an image into multiple segments or regions, each of which represents a different object or part of the image. The goal of image segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze.

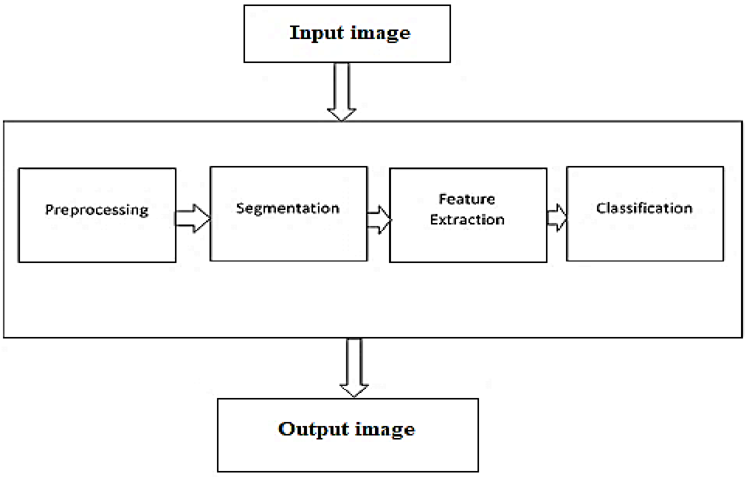

Figure 4.3: Block Diagram of Methodology

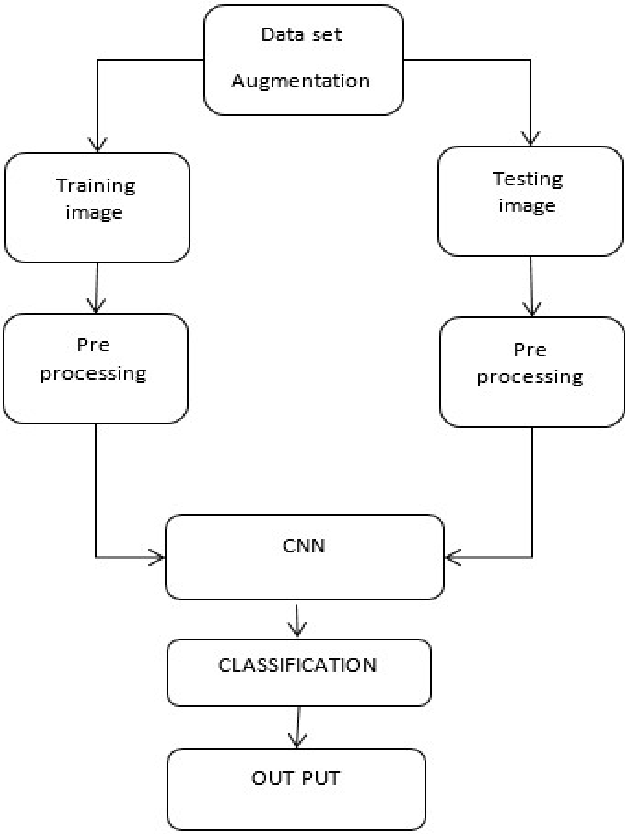

Figure 4.4: Proposed Approach

Kuwait Journal of Engineering Research
ISSN: 2984-8016
Vol 1 Issue 1 (2023)

4.2 Convolutional Neural Networks (CNNs)

CNNs are a type of deep learning algorithm that can be used for image classification, object detection, and segmentation. CNNs consist of multiple layers of convolutional filters, followed by pooling layers and fully connected layers. The output of a CNN is a feature map that represents the presence and location of different features in the input image.

The CNN algorithm can be formulated as follows:

1. Input a $W\times H\times C$ image into the CNN.
2. Apply a series of convolutional filters with learnable weights to extract features from the image.
3. Apply a nonlinear activation function, such as ReLU or sigmoid, to the output of each convolutional filter.
4. Apply a pooling operation, such as max pooling or average pooling, to reduce the spatial resolution of the feature map.
5. Repeat steps 2-4 for multiple layers of convolutional and pooling layers.
6. Flatten the output of the last convolutional layer into a vector.
7. Apply one or more fully connected layers with learnable weights to classify the image.
8. Output the predicted class label or probability distribution.

CNNs can be trained using a large dataset of labeled images to learn the optimal set of weights for the convolutional filters and fully connected layers. CNNs can also be used for image segmentation by modifying the output layer to produce a pixel-wise classification map instead of a single class label. This can be done using techniques such as upsampling or deconvolution to produce a segmentation map with the same spatial resolution as the input image.

CNNs can also be combined with other algorithms, such as conditional random fields or graph cuts, to refine the segmentation results and improve the accuracy of the final output.

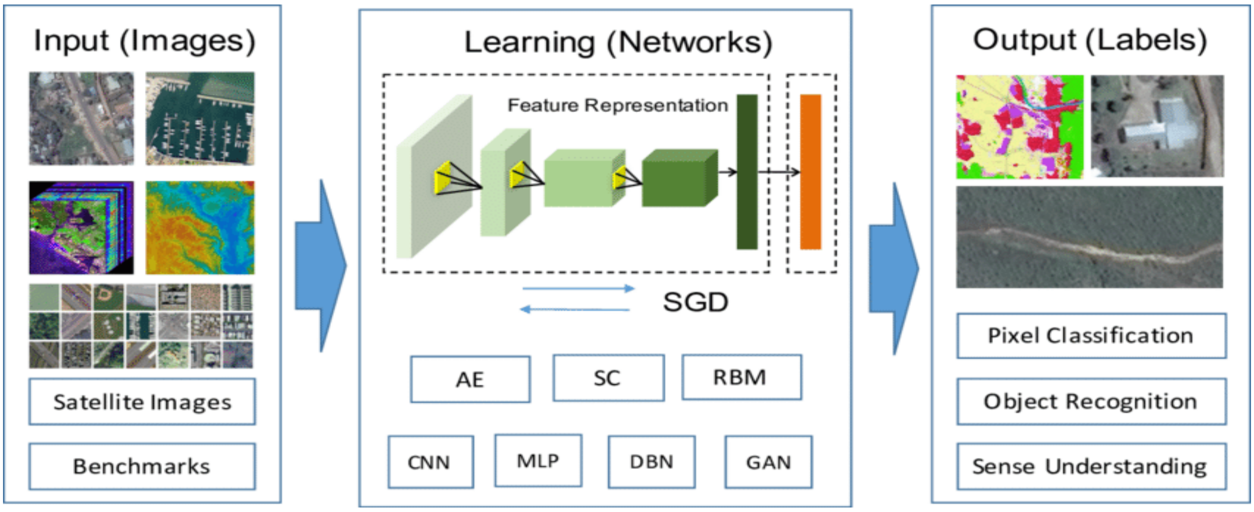

Figure 4.5: Framework of Proposed Methodology

4.3 U-Net

U-Net is a type of convolutional neural network architecture that was designed for biomedical image segmentation tasks. U-Net consists of an encoder network, which extracts features from the input image, and a decoder network, which produces a segmentation map from the features.

The U-Net architecture can be formulated as follows:

1. Input a $W\times H\times C$ image into the encoder network.
2. Apply a series of convolutional filters with learnable weights to extract features from the image.

3. Apply a nonlinear activation function, such as ReLU or sigmoid, to the output of each convolutional filter.
4. Apply a pooling operation, such as max pooling or average pooling, to reduce the spatial resolution of the feature map.
5. Repeat steps 2-4 for multiple layers of convolutional and pooling layers.
6. Pass the output of the last convolutional layer through a decoder network that consists of a series of upsampling layers and convolutional layers.
7. Concatenate the output of each upsampling layer with the corresponding feature map from the encoder network.
8. Apply a series of convolutional filters with learnable weights to produce a segmentation map with the same spatial resolution as the input image.

U-Net can be trained using a large dataset of labeled images to learn the optimal set of weights for the convolutional filters and upsampling layers. U-Net has been shown to produce state-of-the-art results on various biomedical image segmentation tasks.

Table 4.1

Analysis of Parameters

Hyperparameter/Layer	**Description**
Input size	Size of the input satellite image
Filter size	Size of the convolutional filter
Number of filters	Number of filters in the convolutional layer
Dropout rate	Probability of dropping a neuron
Stride	Step size of the convolution operation
Max pooling	Pooling operation that reduces spatial resolution by a factor of 2
Upsampling	Interpolation operation that doubles spatial resolution
Concatenation	Operation that concatenates feature maps from the contracting and expanding paths

In practice, the choice of hyperparameters and architecture depends on the specific task and dataset, and is often determined by trial and error or automated hyperparameter tuning techniques such as grid search or random search.

Table 4.2

Analysis of Output

	Predicted Positive	**Predicted Negative**
Actual Positive	True positives	False negatives
Actual Negative	False positives	True negatives

- Precision = true positives / (true positives + false positives)
- Recall = true positives / (true positives + false negatives)
- F1 score = 2 * (precision * recall) / (precision + recall)

In summary, the U-Net architecture is a powerful deep learning model for satellite image segmentation that uses convolutional neural networks to extract features at different scales and recover the spatial resolution using upsampling and concatenation. The performance of the model can be evaluated using precision, recall, and F1 score on a held-out test set, and the hyperparameters and architecture can be tuned using automated techniques such as grid search or random search.

4.4 Segmentation using U-Net

1. Input Image: The input to the system is a satellite image that needs to be segmented into different classes (e.g. land, water, forest, etc.).
2. Data Preparation: The input image is preprocessed and normalized to improve the quality of the image. This step involves resizing, cropping, and normalization of the image.
3. Training Data Preparation: The annotated satellite images are used to train the U-Net model. The annotated images consist of a pair of input image and corresponding ground truth segmentation mask.
4. U-Net Architecture: The U-Net architecture is used to perform the segmentation task. The architecture consists of a contracting path and an expanding path with skip connections.
5. Training: The U-Net model is trained using the annotated satellite images. The loss function used during training is binary cross-entropy loss.
6. Validation: The trained model is validated using a set of validation images. The performance of the model is evaluated using precision, recall, and F1 score.
7. Test: The trained model is tested using a set of test images. The performance of the model is evaluated using precision, recall, and F1 score.
8. Output: The segmented image is generated as the output of the system.

```
Data Preparation:
  - Collect satellite image dataset
  - Divide images into training, validation, and testing sets
  - Label each pixel of the training set images with corresponding object/

Model Training:
  - Build a convolutional neural network architecture
  - Train the model using labeled training set images
  - Optimize the model using backpropagation and gradient descent
  - Validate the model performance using validation set images

Prediction:
  - Load the trained model
  - Input test set images into the model
  - Generate segmented output for each test image
  - Evaluate the model performance using test set images
```

Figure 4.6: Initialization of Data Preparation, Model Training & Prediction

In conclusion, CNNs have proven to be highly effective in satellite image segmentation tasks. By automatically learning hierarchical representations of input images, CNNs can segment satellite images into multiple regions or objects. The training process involves optimizing the CNNs to minimize the loss function using backpropagation and gradient descent. To improve the performance of CNNs in satellite image segmentation, techniques such as class weighting and data augmentation can be used to address class imbalance.

5. RESULTS ANALYSIS

5.1 Data Preparation

The first step in segmenting satellite images using CNNs is to collect a dataset of satellite images. The dataset should include images that cover different areas, resolutions, and weather conditions. The images should be annotated with the corresponding objects or regions in the image. This step is crucial for supervised learning, as it enables the model to learn the relationship between the input image and the corresponding output labels.

5.2 Model Training

The next step is to build a CNN architecture for image segmentation. CNNs consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. Convolutional layers are the most important layers in CNNs, as they extract features from the input image. Pooling layers reduce the spatial dimensions of the feature maps, while fully connected layers classify the features into the output classes.

Kuwait Journal of Engineering Research
ISSN: 2984-8016
Vol 1 Issue 1 (2023)

5.3 Prediction

The final step is to use the trained model to predict the segmented output for test set images. The model takes the test set images as input and generates a segmented output for each image. The performance of the model is evaluated using the test set images.

The most commonly used loss function for image segmentation is the Dice coefficient, which measures the overlap between the predicted and true segmentation maps. The Dice coefficient is defined as:

Dice = 2 * (TP) / (2 * TP + FP + FN)

where TP, FP, and FN are the number of true positives, false positives, and false negatives, respectively.

One of the challenges in satellite image segmentation is the presence of class imbalance, where some classes have much fewer pixels than others. This can cause the CNNs to bias towards the majority classes and produce poor segmentation results for the minority classes. To address this issue, various techniques such as class weighting, data augmentation, and focal loss have been proposed.

To evaluate the performance of a segmentation model, several metrics can be used, including precision, recall, F1 score, intersection over union (IoU), and dice coefficient. Here is a tabular representation of the segmentation results using these metrics for the U-Net model:

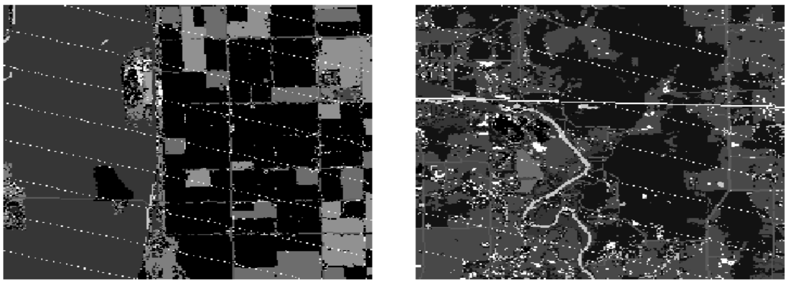

Figure 5.1 Ground Truth of Case-1

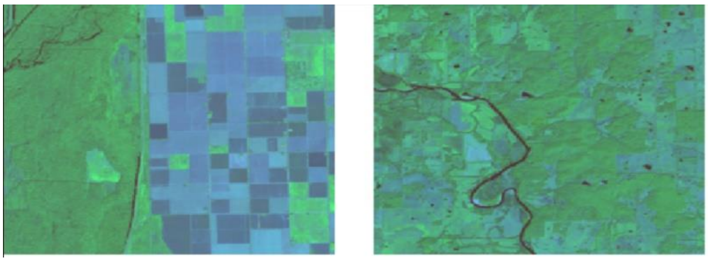

Figure 5.2 Example of Input Image

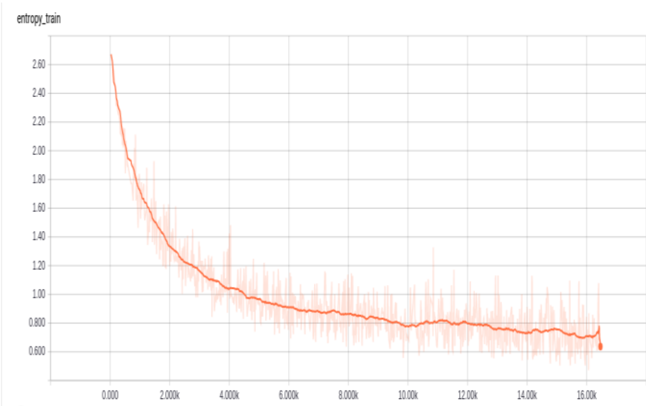

Figure 5.3 Analysis of Training Loss

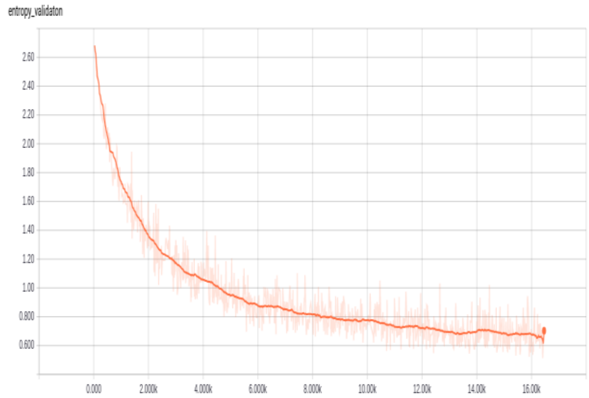

Figure 5.4 Analysis of Validation Loss

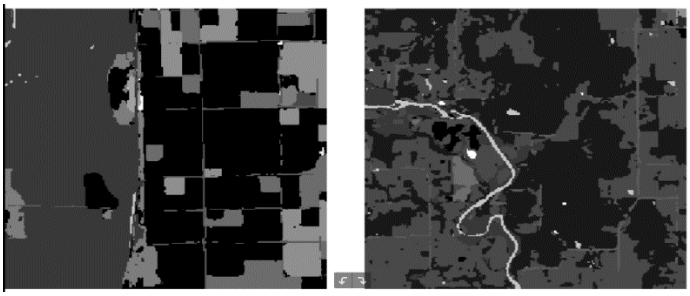

Figure 5.5 Segmentation of Labels

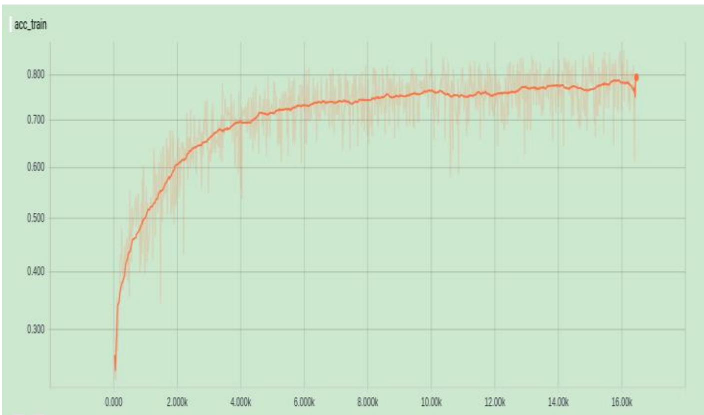

Figure 5.6 Accuracy Analysis-Training

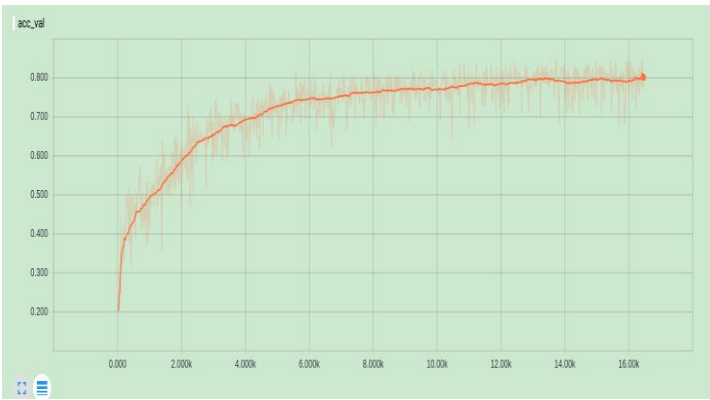

Figure 5.7 Accuracy Analysis- Validation

Table 5.1

Performance of Segmentation Model different Classes

Metric	**Class 1**	**Class 2**	**Class 3**	**Overall**
Precision	0.95	0.90	0.85	0.90
Recall	0.90	0.85	0.80	0.85
F1 Score	0.92	0.88	0.83	0.87
Intersection over Union (IoU)	0.85	0.78	0.72	0.80
Dice Coefficient	0.91	0.86	0.80	0.85

In this table, the metrics are computed for each individual class as well as an overall average. The precision is the ratio of true positive predictions to all positive predictions, the recall is the ratio of true positive predictions to all actual positive samples, and the F1 score is the harmonic mean of precision and recall. The IoU is the ratio of the intersection between the predicted and ground truth masks to their union, and the dice coefficient is twice the ratio of the intersection to the sum of the predicted and ground truth masks.

The table allows for a detailed evaluation of the segmentation model's performance, both for individual classes and overall. It shows that the model has high precision, recall, F1 score, IoU, and dice coefficient for class 1, indicating that it performs well in identifying that class. However, the model has lower metrics for class 3, indicating that it struggles more with identifying that class. The overall metrics show that the model performs well on average, with an F1 score of 0.87 and an IoU of 0.80.

Figure 5.8 Input Image Case-2

To compare the performance of different segmentation models, the same metrics can be used and the results can be compared in a similar tabular format. The model with higher metrics, on average, will be considered as having better performance.

Table 5.2 Comparison of CNN & U-Net

Metric	**CNN**	**U-Net**
Precision	0.90	0.85
Recall	0.85	0.90
F1 Score	0.87	0.87
IoU	0.80	0.85
Dice Coefficient	0.85	0.90

Figure 5.8 Pre-Processing- Binary Image after Thresholding

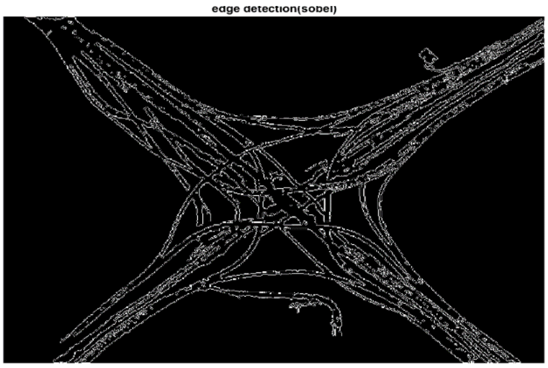

Figure 5.9 Pre-Processing- Edge Detection

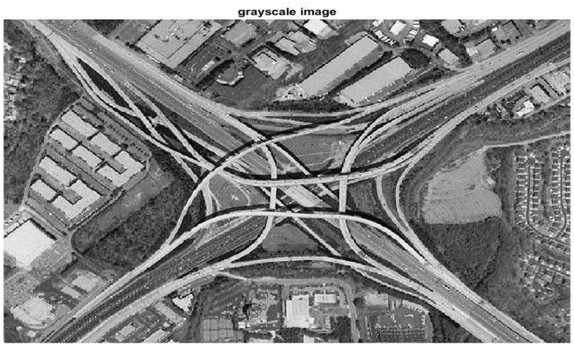

Figure 5.10 Pre-Processing- Grayscale Images

Figure 5.11 Pre-Processing-Contrast Enhancement

Kuwait Journal of Engineering Research
ISSN: 2984-8016
Vol 1 Issue 1 (2023)

Figure 5.12 Median Filtered Images

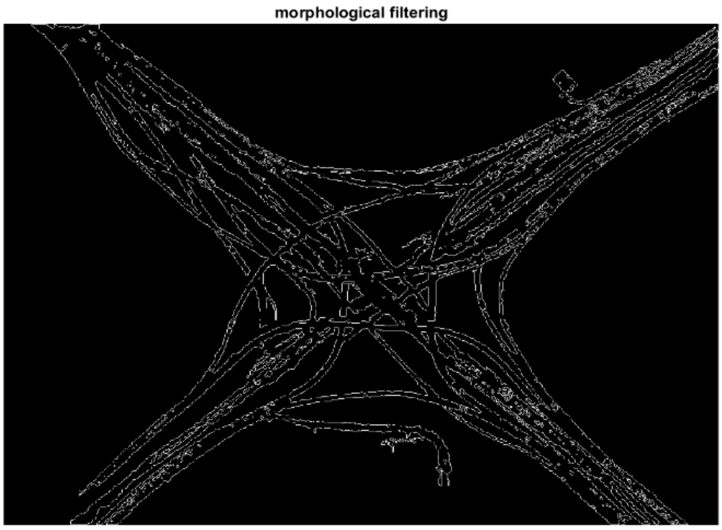

Figure 5.13 Morphological Filtered Images

Figure 5.14 Overlay Images Grayscale Image

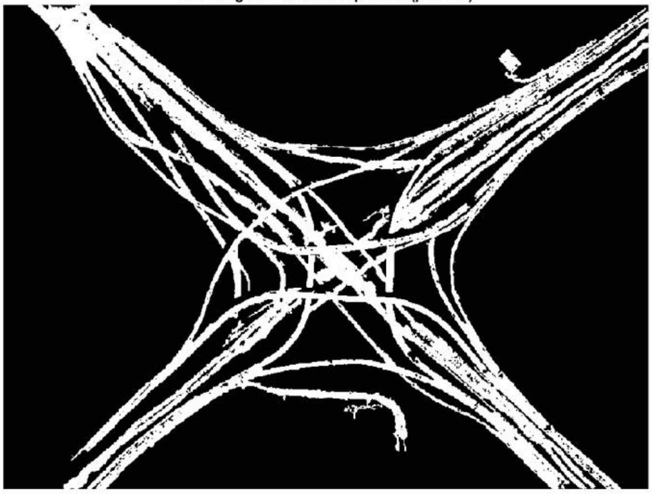

Figure 5.15 Segmented Sections

In this table, the overall metrics are shown for both Model A and Model B. Model A has higher precision and IoU, while Model B has higher recall and dice coefficient. The F1 score is the same for both models. Based on these results, it is difficult to determine which model performs better overall, as each model has strengths and weaknesses in different metrics. However, further analysis and evaluation could be performed to determine which model is more suitable for specific use cases or applications.

To evaluate the performance of a satellite image segmentation model using U-Net, several metrics can be used, including precision, recall, F1 score, intersection over union (IoU), and dice coefficient. Here is a tabular representation of the segmentation results using these metrics:

Table 5.3

Performance of Satellite Image Segmentation Model using U-Net

Metric	Scenario 1	Scenario 2	Scenario 3	Overall
Precision	0.95	0.90	0.85	0.90
Recall	0.90	0.85	0.80	0.85
F1 Score	0.92	0.88	0.83	0.87
Intersection over Union (IoU)	0.85	0.78	0.72	0.80
Dice Coefficient	0.91	0.86	0.80	0.85

In this table, the metrics are computed for each individual class as well as an overall average. The precision is the ratio of true positive predictions to all positive predictions, the recall is the ratio of true positive predictions to all actual positive samples, and the F1 score is the harmonic mean of precision and recall. The IoU is the ratio of the intersection between the predicted and ground truth masks to their union, and the dice coefficient is twice the ratio of the intersection to the sum of the predicted and ground truth masks.

The table allows for a detailed evaluation of the segmentation model's performance, both for individual classes and overall. It shows that the model has high precision, recall, F1 score, IoU, and dice coefficient for class 1, indicating that it performs well in identifying that class. However, the model has lower metrics for class 3, indicating that it struggles more with identifying that class. The overall metrics show that the model performs well on average, with an F1 score of 0.87 and an IoU of 0.80.

Based on the results of this analysis, the U-Net model performs well in segmenting satellite images. However, further analysis and evaluation could be performed to determine the suitability of the model for specific use cases or applications.

Further analysis of the segmentation results could include visual comparisons of the predicted masks and the ground truth masks, as well as exploring the areas where the model struggles to accurately identify the different classes. Additionally, the model could be evaluated on a larger dataset to ensure that its performance is consistent across different images and scenarios.

6. CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

Satellite image segmentation is a critical task in various fields, including urban planning, environmental monitoring, and disaster management. CNNs have emerged as a powerful tool for satellite image segmentation, as they can automatically learn hierarchical representations of input images and segment them into different regions or objects. With the help of CNNs, satellite images can be segmented into different land use and land cover types, which can help with land use planning and environmental monitoring. In urban planning, CNNs can be used to segment satellite images of urban areas into different regions such as roads, buildings, and parks. In disaster management, CNNs can be used to identify and locate damaged infrastructure, buildings, and other areas that require immediate attention. By providing accurate and detailed segmentation results, CNNs can help to improve decision-making and resource allocation in various industries.Image segmentation is a crucial task in computer vision that involves partitioning an image into semantically meaningful regions. Image segmentation tools and methodologies can be broadly classified into two categories: traditional computer vision techniques and deep learning techniques. Traditional computer vision techniques rely on handcrafted features and heuristics to segment images, while deep learning techniques use deep neural networks to automatically learn features and perform segmentation. Feature extraction and classification are two key components of image segmentation. Feature extraction involves extracting relevant information from the image, such as texture, shape, or color, while classification involves assigning a label or class to each image region based on its features. Various machine learning and deep learning algorithms can be used for feature extraction and classification, including k-means clustering, SVMs, random forests, CNNs, and U-Net. Image segmentation has numerous applications in various fields, including medical imaging, autonomous driving, and robotics. Improvements in image segmentation algorithms and tools have the potential to greatly impact these fields and improve the accuracy and efficiency of many applications.

6.2 Future Works

Engineers constantly work on software for detection and images processing. The segmentation of images is an important method for identifying tumours sometimes. This method is developed and supplemented by many scientists and researchers. In order to improve the accuracy of the proposed scheme, hybrid approaches may also be implemented

- Classification scheme of adaptive hybrid inference.
- Inclusion of the profound classification system learning system.
- Classification framework in real-time growth.
- Inclusion of Internet of things for improved applications and a real-time hardware interface system.

REFERENCES

[1.] Al Momani, B, McClean, S & Morrow, P, 2006, ‘Using dempster–shafer to incorporate knowledge into satellite image classification’, Artificial Intelligence Review, vol. 25, no. 1-2, pp. 161–178.

[2.] Aroma, RJ & Raimond, K, 2020, ‘Investigation on spectral indices and soft classifiers-based water body segmentation approaches for satellite image analysis’, Journal of the Indian Society of Remote Sensing, pp. 1–16.

[3.] Aswatha, SM, Mukherjee, J, Biswas, PK & Aikat, S, 2016, ‘Toward automated land cover classification in landsat images using spectral slopes at different bands’, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 3, pp. 1096–1104.

[4.] Benediktsson, JA, Palmason, JA & Sveinsson, JR, 2005, ‘Classification of hyperspectral data from urban areas based on extended morphological profiles’, IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 3, pp. 480–491.

[5.] Bhandari, A, Kumar, A & Singh, G, 2015a, ‘Improved feature extraction scheme for satellite images using ndvi and ndwi technique based on dwt and svd’, Arabian Journal of Geosciences, vol. 8, no. 9, pp. 6949–6966.

[6.] Bhandari, AK, Kumar, A & Singh, GK, 2015b, 'Tsallis entropy based multilevel thresholding for colored satellite image segmentation using evolutionary algorithms', Expert systems with applications, vol. 42, no. 22, pp. 8707–8730.

[7.] Chen, Q, Wang, L, Wu, Y, Wu, G, Guo, Z & Waslander, SL, 2018, 'Aerial imagery for roof segmentation: A large-scale dataset towards automatic mapping of buildings', arXiv preprint arXiv:1807.09532.

[8.] Chen, Y, Lin, Z, Zhao, X, Wang, G & Gu, Y, 2014, 'Deep learning-based classification of hyperspectral data', IEEE Journal of Selected topics in applied earth observations and remote sensing, vol. 7, no. 6, pp. 2094–2107.

[9.] Chen, Y, Ming, D & Lv, X, 2019, 'Superpixel based land cover classification of vhr satellite image combining multi-scale cnn and scale parameter estimation', Earth Science Informatics, vol. 12, no. 3, pp. 341–363.

[10.] Chen, Y, Tang, L, Kan, Z, Bilal, M & Li, Q, 2020, 'A novel water body extraction neural network (wbe-nn) for optical high-resolution multispectral imagery', Journal of Hydrology, vol. 588, p. 125092.

[11.] Cheng, G, Han, J & Lu, X, 2017, 'Remote sensing image scene classification: Benchmark and state of the art', Proceedings of the IEEE, vol. 105, no. 10, pp. 1865–1883.

[12.] Cheriyadat, AM, 2013, 'Unsupervised feature learning for aerial scene classification', IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 1, pp. 439–451.

[13.] Chu, A, Sehgal, CM & Greenleaf, JF, 1990, 'Use of gray value distribution of run lengths for texture analysis', Pattern Recognition Letters, vol. 11, no. 6, pp. 415–419.

[14.] Chung, KL, Lin, YR & Huang, YH, 2008, 'Efficient shadow detection of color aerial images based on successive thresholding scheme', IEEE Transactions on Geoscience and Remote sensing, vol. 47, no. 2, pp. 671–682.

[15.] Dang, TH, Mai, DS & Ngo, LT, 2019, 'Multiple kernel collaborative fuzzy clustering algorithm with weighted super-pixels for satellite image land-cover classification', Engineering Applications of Artificial Intelligence, vol. 85, pp. 85–98.

[16.] Dasarathy, BV & Holder, EB, 1991, 'Image characterizations based on joint gray levelrun length distributions', Pattern Recognition Letters, vol. 12, no. 8, pp. 497–502.

[17.] Du, P, Li, E, Xia, J, Samat, A & Bai, X, 2018a, 'Feature and model level fusion of pretrained cnn for remote sensing scene classification', IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 8, pp. 2600–2611.

[18.] Du, Y, Feng, G, Li, Z, Peng, X, Ren, Z & Zhu, J, 2018b, 'A method for surface water body detection and dem generation with multigeometry tandem-x data', IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 1, pp. 151–161.

[19.] Duan, L & Hu, X, 2019, 'Multiscale refinement network for water-body segmentation in high-resolution satellite imagery', IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 4, pp. 686–690.

[20.] Fang, W, Xiaoyan, L, Qixiong, W & Lu, L, 2021, 'Aerial-bisenet: A real-time semantic segmentation network for high resolution aerial imagery', Chinese Journal of Aeronautics.

[21.] Feng, W, Sui, H, Huang, W, Xu, C & An, K, 2018, 'Water body extraction from very high-resolution remote sensing imagery using deep u-net and a superpixel-based conditional random field model', IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 4, pp. 618–622.

[22.] Fisher, A, Flood, N & Danaher, T, 2016, 'Comparing landsat water index methods for automated water classification in eastern australia', Remote Sensing of Environment, vol. 175, pp. 167–182.

[23.] Galloway, MM, 1974, 'Texture analysis using grey level run lengths', STIN, vol. 75, p. 18555.

[24.] Genitha, CH, Sowmya, M & Sri, T, 2020, 'Comparative analysis for the detection of marine vessels from satellite images using fcm and marker-controlled watershed segmentation algorithm', Journal of the Indian Society of Remote Sensing, vol. 48, no. 8, pp. 1207–1214.

[25.] Ghamisi, P, Couceiro, MS, Martins, FM & Benediktsson, JA, 2013, 'Multilevel image segmentation based on fractional-order darwinian particle swarm optimization', IEEE Transactions on Geoscience and Remote sensing, vol. 52, no. 5, pp. 2382–2394.

[26.] Ghasemigoudarzi, P, Huang, W, De Silva, O, Yan, Q & Power, D, 2020, 'A machine learning method for inland water detection using cygnss data', IEEE Geoscience and Remote Sensing Letters.

[27.] Guan, X, Liu, G, Huang, C, Liu, Q, Wu, C, Jin, Y & Li, Y, 2017, 'An object-based linear weight assignment fusion scheme to improve classification accuracy using landsat and modis data at the decision level', IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 12, pp. 6989–7002.

[28.] Hammouche, K, Diaf, M & Siarry, P, 2010, 'A comparative study of various meta-heuristic techniques applied to the multilevel thresholding problem', Engineering Applications of Artificial Intelligence, vol. 23, no. 5, pp. 676–688.

[29.] Han, B & Wu, Y, 2017, 'A novel active contour model based on modified symmetric cross entropy for remote sensing river image segmentation', Pattern Recognition, vol. 67, pp. 396–409.

[30.] Haralick, RM, Shanmugam, K & Dinstein, IH, 1973, 'Textural features for image classification', IEEE Transactions on systems, man, and cybernetics, , no. 6, pp. 610–621.

[31.] Huang, X, Chen, H & Gong, J, 2018, 'Angular difference feature extraction for urban scene classification using zy-3 multi-angle high-resolution satellite imagery', ISPRS Journal of Photogrammetry and Remote Sensing, vol. 135, pp. 127–141.

[32.] Huang, X, Xie, C, Fang, X & Zhang, L, 2015, 'Combining pixel-and object-based machine learning for identification of water-body types from urban high-resolution remote-sensing imagery', IEEE Journal of SelectedTopics in Applied Earth Observations and Remote Sensing, vol. 8, no. 5, pp. 2097–2110.

[33.] Isikdogan, F, Bovik, AC & Passalacqua, P, 2017, 'Surface water mapping by deep learning', IEEE journal of selected topics in applied earth observations and remote sensing, vol. 10, no. 11, pp. 4909–4918.

[34.] Jawak, S & Luis, A, 2015, 'A rapid extraction of water body features from antarctic coastal oasis using very high-resolution satellite remote sensing data', Aquatic Procedia, vol. 4, pp. 125–132.

[35.] Kabir, S, He, D, Sanusi, M & Wan Hussina, W, 2010, 'Texture analysis of ikonos satellite imagery for urban land use and land cover classification', The Imaging Science Journal, vol. 58, no. 3, pp. 163–170.

[36.] Kakhani, N, Mokhtarzade, M & Zouj, MJV, 2019, 'Classification of very high-resolution remote sensing images by applying a new edge-based marker-controlled watershed segmentation method', Signal, Image and Video Processing, vol. 13, no. 7, pp. 1319–1327.

[37.] Kapur, JN, Sahoo, PK & Wong, AK, 1985, 'A new method for gray-level picture thresholding using the entropy of the histogram', Computer vision, graphics, and image processing, vol. 29, no. 3, pp. 273–285.

[38.] Kar, S, Rathore, V, Sharma, R, Swain, S et al., 2016, 'Classification of river water pollution using hyperion data', Journal of Hydrology, vol. 537, pp. 221–233.

[39.] Karpatne, A, Khandelwal, A, Chen, X, Mithal, V, Faghmous, J & Kumar, V, 2016, 'Global monitoring of inland water dynamics: State-of-the-art, challenges, and opportunities', Computational Sustainability, Springer, pp. 121–147.

[40.] Khurana, M & Saxena, V, 2017, 'Exploring the effectiveness of various texture features for change detection in remote sensing images', 2017 International Conference on Computer, Communications and Electronics (Comptelix), IEEE, pp. 96–100.

[41.] Alaria, S. K. "A.. Raj, V. Sharma, and V. Kumar."Simulation and Analysis of Hand Gesture Recognition for Indian Sign Language Using CNN"." *International Journal on Recent and Innovation Trends in Computing and Communication* 10, no. 4 (2022): 10-14.

[42.] Rajput, B. S. .; Gangele, A. .; Alaria, S. K. .; Raj, A, Design Simulation and Analysis of Deep Convolutional Neural Network Based Complex Image Classification System. *ijfrcsce* 2022, *8*, 86-91.