Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

ENHANCING LAND USE AND LAND COVER ANALYSIS THROUGH IMAGE SEGMENTATION TECHNIQUES

SMITA SUNIL BURREWAR *
Research Scholar, Department of Architecture and Planning, National Institute of Technology, Patna, India.
*Corresponding Author Email: smitab.phd20.ar@nitp.ac.in

MAZHARUL HAQUE
Assistant Professor, Department of Architecture and Planning, National Institute of Technology, Patna, India. Email: mazharul@nitp.ac.in

TANWIR UDDIN HAIDER
Associate Professor, Department of Computer Science & Engineering, National Technology of Engineering, India. Email: tanwir@nitp.ac.in

Abstract

Land use and land cover (LULC) analysis is a fundamental component of environmental monitoring and land management, offering valuable insights for urban planning and sustainable development. With the advancements in machine learning, new avenues have opened for gaining insights about urbanization trends, deforestation and climate change. There has been works done on LULC but they have used aerial images directly without applying image segmentation resulting in limited insights, inaccuracies, lack of localization and inability to distinguish between complex land cover types. To solve the drawbacks, we have applied image segmentation techniques to aerial images of LULC in this study. Five types of image segmentation techniques are used: 1) Threshold based, 2) Edge-based, 3) K-means clustering based, 4) Otsu's segmentation, 5) Unet with ResUnet. The aerial images are divided into ten types of land uses. Dataset was prepared by acquiring aerial images of LULC, the images were pre-processed to enhance its quality and suitability for segmentation, and finally then one by one each segmentation technique was applied to all the images. Through experimentation and validation the most suitable image segmentation technique for LULC has been determined. The Accuracy, Jaccard Similarity Coefficient (JSC), Dice Similarity Coefficient (DSC), Intersection Over Union (IoU), Mean Intersection Over Union (MIOU) have been used to assess the effectiveness on LULC segmentation.

Keywords: Land Use Land Cover Changes, Image Segmentation, Resunet, Urbanization, And Environmental Monitoring.

1. INTRODUCTION

The impact of changes in land use and cover (LULC) on ecological processes has received more attention in recent years. Monitoring these variations can be useful for planning in both urban and rural regions, determining seasonal variations in temperature, and monitoring the environment. The most significant factor influencing the global distribution of climate and biogeochemistry is natural variations in land cover. Strength balance, biogeochemical cycles, and the hydrologic cycle are all significantly impacted by the Earth's surface. Consequently, these cycles are greatly impacted by the Earth's surface, which also has an impact on the balance of strength. Climate change results from human-caused actions such overgrazing, baring, and trapping, as well as from human-caused activities to a lesser extent (Smita.S.Burrewar, 2024).

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

The surface of the planet is changing gradually, yet regular processes like morphological evolution, vegetation type shifts, and temperature change caused by orbital mechanics are what cause some of these variations in land cover. The scientific community understands the need of mapping, observing, and assessing the impacts of modifications to the physical features of the Earth's surface (Shi et al., 2021). The manner, in which individuals make use of their land for commercial purposes, as well as for other social and cultural purposes including recreation and preservation, is referred to as land use (LU). The percentage of Earth's land area that is covered by different kinds of flora, water, forests, and other natural features is referred to as land cover (LC). The world's population increase in the modern era has caused severe challenges for society and the environment, as evidenced by the growing need for resources, services, food, shelter, and other necessities. Changes in land use and land cover (LULC) have resulted from these variables, causing adverse impacts on the ecosystem. In addition to undoing the negative effects of earlier LU decisions, effective land use planning and management are becoming increasingly crucial to the long-term sustainability and well-being of emerging nations. Land-use sceneries include port, house, highway, farming, and park, since each one uses a variety of features to depict a distinct situation (Cao et al., 2019). Due to advancement in several optical satellite sensors that can capture high-resolution images, studies of land cover using these kinds of images have become quite popular in the field of remote sensing (Wang et al., 2020). The capacity to identify changes on the surface of the Earth facilitates our understanding of the interactions between humans and their natural settings (Shi et al., 2021). Since the land resource provides the physical basis for social activity and economic growth, it is essential to the development of a region's economy, society, and ecological health overall. As such, it must be considered in any discussion on environmental sustainability.

Sustainable design is based on four main principles: place, natural processes, people, and the environment's influence. Using co-creative design techniques also facilitates connecting with nature. Design practices are influenced by place; the impact of management and economics on the environment is determined by knowledge of natural processes; the diversity of cultural practices and human behavior is taken into account by knowledge of people; and design outcomes are enhanced by knowledge of management and economics. The construction process can lead to the indirect or direct release of industrial effluent, waste gas, and waste residue, which can contaminate the soil and cause acidification, salinization, and hardness of the soil (Liu et al. 2020). Being a dynamic process with significant environmental implications, many techniques have been developed to evaluate and simulate land use change (LUC) for sustainable land use management and decision-making. The stresses that land use change (LUC) as a result of urban growth places on the ecosystem might have irreversible effects if left unchecked (Shi et al., 2021). By recognizing the characteristics of land cover and reflecting changes in local human behavior and socioeconomic conditions in an area, land use classification offers valuable information for managing and planning land resources, anticipating and monitoring urban growth, and preserving natural systems. Remotely sensed images are frequently used for land-use classification due to the abundance of structural and spatial

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

information they offer about ground objects. Land-use categorization, a popular field of research in remote sensing, involves giving extracted areas that contain several land-use classifications semantic labels. In order to develop robust, powerful, and discriminative features, deep learning models are currently commonly utilized in feature learning due to their excellent feature representation capabilities in a number of applications.

Using aerial photos directly without image segmentation indeed poses several limitations in land use and land cover (LULC) studies. Image segmentation is a crucial step in remote sensing and image processing that divides an image into meaningful segments or regions. The application of image segmentation in LULC studies is essential for achieving accurate localization, gaining deeper insights, discerning between complicated land cover types, improving classification accuracy, and facilitating spatial analysis. Integrating segmentation techniques into the analysis pipeline enhances the effectiveness of aerial photo interpretation and contributes to more informed decision-making in land management and environmental conservation efforts. Image segmentation techniques have revolutionized the analysis, interpretation, and use of visual data, offering several benefits in a variety of sectors. The primary benefit is their capacity to offer accurate object localization in photos. These approaches allow precise identification and delineation of certain elements, such as land cover types in satellite imagery and tumors in medical scans, by segmenting images into separate sections. This accuracy not only helps us comprehend complicated scenes better, but it also makes focused analysis and decision-making easier. Moreover, segmentation makes it easier to extract features, which makes it possible to extract important characteristics like color, texture, and shape. With this information collected, processes like analysis, classification, and interpretation may be carried out more easily and effectively, enabling a wide range of applications from autonomous driving to LULC analysis. Images of various land usage is presented in Fig 1.**Error! Reference source not found.**

Fig 1: Aerial Images showing various land usage

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

Segmentation also helps with semantic image annotation, which makes it possible to automatically name and classify images according to their content. Segmentation facilitates region-based analysis in addition to classification, allowing researchers to examine spatial patterns, distributional features, and connections between various parts within an image. All things considered, image segmentation methods constitute a cornerstone of image analysis, offering priceless instruments for extracting the abundance of information contained in visual data and spurring creativity in a wide range of fields.

A diverse range of land use forms have quickly replaced the world's land cover in recent decades. When describing global environmental transitions, the words "changes in land use" and "changes in land cover," which are distinct but related, are frequently used interchangeably. It is essential to comprehend LULC and its impacts for long-term planning and efficient resource management. Technical, social, and institutional frameworks work together to determine land use and cover, taking into account ecological parameters, altitudes, and ecological structure (Mishra et al., 2020). The human-environment interaction that gives birth to land use change (LUC) is still in its infancy, and as a result, around 39% of Earth's land has never been utilized by people in any way (Cao et al., 2019).

2. RELATED WORKS

Numerous research studies, from machine learning to deep learning, have examined land cover. Ten years ago, neural networks' enormous computational complexity made their utilisation undesirable. While histogram thresholding yielded satisfactory results, it displayed issues related to the inconsistencies and difficulties in satellite imagery. In a similar vein, LULC mapping employed traditional machine learning algorithms like random forest techniques and support vector machines. For instance, these techniques are used in (Bengana & Heikkilä, 2021), (Tian, Li, & Shi, 2018) to classify land cover. A decision tree and an artificial neural network were used to categorise land cover using Landsat ETM+ data in the machine learning land cover classification project. The disadvantage of these approaches is that, in order to enhance model performance, a thorough understanding of the feature extraction procedure is necessary.

Nonetheless, current research indicates that segmentation and classification tasks frequently employ deep learning methods. Deep learning performs particularly well in large-unity language classification (LULC) tasks including building detection, urbanisation, water resource monitoring, agricultural field monitoring, and forest change detection because satellite images include a higher number of features and a more complicated structure. In order to construct a satellite image with building masks, for instance, automatic building recognition algorithms were used in (Chhor, Aramburu, & Bougdal-Lambert, 2017), where a dataset was gathered via the Map Box API for Open Street Map. Additionally, study (Karwowska & Wierzbicki, 2022) explains the pixel-by-pixel image segmentation techniques for categorising various satellite picture properties. Here, the suggested approach can identify a building in the INRIA dataset—which

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

consists of incredibly high-resolution photos—with a high degree of accuracy by utilising the UNet model. These research, meanwhile, were limited to class segmentation. It is more difficult to develop a model that makes advantage of multi-class segmentation of satellite images. While ResNet, VGG, and EfficientNet are utilised for computer vision classification problems, deep learning architectures like UNet and DenseNet are being used for image segmentation. Recent studies' findings indicate that these deep learning models perform better than traditional feature extraction techniques.

To attain great performance, however, further work needs to be done in the area of satellite image processing. For instance, because of the intricate structure of satellite photos, the outcomes of contemporary semantic segmentation were unsatisfactory (Wafa, Khan, Malik, Abdusalomov, Cho, & Odarchenko, 2022), (Abdusalomov, Mukhiddinov, Djuraev, Khamdamov, & Whangbo, 2020), (Sevak, Kapadia, Chavda, Shah, & Rahevar, 2017), (Huang, Liu, & Weinberger, 2017). One of the best outcomes in the DeepGlobe challenge was achieved by Kuo et al.'s (Kuo, Tseng, Yan, Liu, & Wang, 2018) approach, which relies on a DeepLabV3+ modification to improve model performance. Despite this, the fixed value of the standard deviation Gaussian filter contributes to their poor model accuracy. Using DeepLab v3+ and the DeepGlobe dataset, Renee Su et al. suggested a semantic segmentation model with an IoU score of 0.756 (Lee, Park, Son, Han, Kim, & Kim, 2019). But because the scientists did not employ any augmentation techniques, their model needs more satellite photos to train. Among the many picture segmentation methods, SegNet, deep convolutional encoder-decoder architecture, is a very effective model. In order to classify the land cover in an aerial image, Lee et al. used the SegNet model. They then conducted research to see how accurate the classification was (Lee, Park, Son, Han, Kim, & Kim, 2019).

Using DeepLab and ResNet18 as the foundation, the authors of (Demir, et al., 2018) presented an architecture that achieved an IoU score of 0.433 s for the DeepGlobe land cover data. Two neural network designs were used by the authors of this study's transfer learning approach. The categorization was done using the ResNet50 model. A pre-trained ResNet50 model was employed as an encoder in the modified UNet model for segmentation following classification (Ulmas & Liiv, 2020). The authors assert that the low accuracy can be attributed mostly to the dataset's quality. The authors also come to the conclusion that machine learning algorithms cannot be trained on the CORINE dataset.

Agricultural field monitoring is one of the key elements of LULC. Using low resolution photos, several research were carried out for agricultural segmentation (Sharifzadeh, Tata, Sharifzadeh, & Tan, 2019), (Kutlimuratov, Abdusalomov, & Whangbo, 2020). Nevertheless, in (Sertel, Ekim, Osgouei, & Kabadayi, 2022), scientists used VHR Worldview-3 photos to create a new benchmark dataset for twelve different LULC classes across two separate geographic areas. While high-resolution images should be utilised for segmenting individual things, such as multi-class segmentation and small objects, low-resolution satellite photos can be used for segmentation jobs of broad or general changes in areas.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

3. BACKGROUND STUDY

Land Use and Land Cover (LULC) analysis plays a crucial role in understanding and managing the Earth's surface and its changes over time. Traditionally, researchers have relied on aerial images for LULC analysis, but there have been limitations in the approach. Without proper segmentation, it's challenging to extract meaningful information from aerial images. Researchers may struggle to identify distinct features or patterns within the images, limiting the depth of analysis. Analyzing aerial images without segmentation can lead to inaccuracies in land cover classification. Without delineating boundaries accurately, there's a higher chance of misclassifying different land cover types. Aerial images cover large areas, and without segmentation, it's difficult to localize specific features or areas of interest within the images. This lack of localization can hinder targeted analysis or monitoring efforts. Complex land cover types, such as mixed-use areas or transitional zones, can be challenging to differentiate without segmentation. This limitation can affect the accuracy of land cover classification and change detection.

Image segmentation involves partitioning an image into multiple segments or regions based on certain characteristics, such as color, texture, or intensity. By segmenting aerial images before analysis, researchers can achieve several benefits like; Image segmentation helps extract more meaningful information from aerial images by identifying coherent regions or objects within the images. This allows researchers to better understand spatial patterns and relationships in land cover distribution. Segmentation enables more accurate delineation of land cover boundaries, reducing the likelihood of misclassification errors. By partitioning the image into homogeneous regions, researchers can better classify different land cover types. Segmenting aerial images facilitates the localization of specific features or areas of interest within the images. This localization capability is essential for targeted analysis, monitoring, and decision-making in land management. Image segmentation helps distinguish complex land cover types by identifying subtle variations in spectral or spatial characteristics. This capability improves the accuracy of classifying mixed or transitional land cover categories.

4. PROBLEM FORMULATION

This paper uses Unet with ResUnet as a backbone model for semantic segmentation of images. Different image segmentation techniques such as: Edge-based segmentation, Threshold based segmentation, K-means Clustering based segmentation, Otsu's segmentation and Artificial Neural Network (ANN) has been applied to aerial images of Land use and Land cover. The aerial images are divided into ten types of land uses according to the features available in them. Five different types of image segmentation techniques have been applied to each type of land use to find out the most suitable image segmentation technique for LULC. The accuracy evaluation framework would be strengthened by the development and application of a fully generalizable technique. It is also crucial to recognize and replicate LC and LU changes over time using the JDL (Joint Deep Learning) framework. The potential of ANNs, which are increasingly being used in land-use categorization assignments, has to be further investigated. This study's primary

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

objective is to find out the best image segmentation technique for Land use land cover analysis.

5. RESEARCH OBJECTIVE

- To apply five different types of image segmentation techniques to LULC aerial images and find out the best segmentation technique.
- For the best results, it is important that all the images are pre-processed so that the noise levels are minimal.

6. PROPOSED METHODOLOGY

Step 1: Data collection and acquisition

In this study, the dataset is acquired from Sentinel-2 satellite. The dataset contains ten distinct classes that fulfill the need for a more comprehensive dataset containing land-use land cover characteristics from various locations and provinces. Each image has a ground sampling distance of 10 meters and a size of 64x64 pixels.

Step 2: Data pre-processing

Pre-processing of aerial images, particularly for segmentation tasks, typically involves several steps to enhance the quality of the data and improve the performance of subsequent segmentation algorithms.

(a) Hyper spectral Image Stack:

- ***Dimensionality Reduction*** - Hyper spectral images often have a high number of spectral bands, which can lead to computational challenges and redundancy. Techniques like Principal Component Analysis (PCA) or t-distributed Stochastic Neighbor Embedding (t-SNE) may be employed to reduce the dimensionality of the hyper spectral data while preserving important spectral information.
- ***Normalization*** - Normalize the pixel values across different bands to ensure consistency in the range of values, which helps in preventing any particular band from dominating the segmentation process.
- ***Feature Extraction*** - Extract relevant features from the hyper spectral image stack that capture unique spectral signatures of different land cover types. This may involve techniques such as spectral indices (e.g., NDVI, NDWI) or texture analysis to highlight specific characteristics.

(b) Atmospheric Correction:

- ***Radiometric Correction*** - Correct for variations in the brightness and contrast of the image caused by atmospheric scattering and absorption. This involves applying algorithms to remove atmospheric effects and normalize radiance values across the image.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

- ***Topographic Correction*** - Correct for variations in illumination caused by terrain elevation, especially in hilly or mountainous regions. Algorithms such as Minnaert correction or C-correction can be applied to standardize illumination conditions across the image.
- ***Atmospheric Compensation*** - Compensate for atmospheric effects such as haze or fog, which can obscure details in the image. This may involve using image enhancement techniques like histogram equalization or dehazing algorithms to improve visibility and contrast.

(c) Additional Pre-processing steps:

- ***Image Registration*** - Align multiple images from different sensors or time periods to ensure spatial consistency, especially in cases where data from multiple sources are combined.
- ***Noise Reduction*** - Apply filters such as median filtering or Gaussian smoothing to reduce noise and improve the signal-to-noise ratio of the image.
- ***Resampling*** - Adjust the spatial resolution of the image to match the requirements of the segmentation algorithm or to ensure consistency with other datasets.

Step 3: Application of Image Segmentation techniques

In this step 3, different image segmentation techniques have been applied;

(a) ***Threshold based segmentation*** - Threshold-based segmentation involves partitioning an image based on pixel intensity values. Pixels with intensities above or below a certain threshold are assigned to different segments. A threshold value is selected based on the histogram of pixel intensities. Pixels are classified as foreground or background depending on whether their intensity values are above or below the threshold.

(b) ***Edge-based segmentation*** - Edge-based segmentation involves detecting edges in an image, which represent abrupt changes in intensity or color. These edges often correspond to object boundaries or significant features in the image. Techniques such as gradient-based edge detection (e.g., Sobel, Prewitt, and Canny) are commonly used to identify edges. Once edges are detected, they can be used as boundaries to segment the image into regions.

(c) ***K-means Clustering based segmentation*** - K-means clustering is an unsupervised machine learning algorithm used for clustering data points into K clusters. In image segmentation, K-means is applied to group similar pixels into clusters. Initially, K cluster centers are randomly initialized. Then, pixels are assigned to the nearest cluster centroid based on pixel intensity or feature vectors. Cluster centroids are updated iteratively until convergence.

(d) ***Otsu's segmentation*** - Otsu's method is a thresholding technique used to automatically determine the optimal threshold value for image segmentation. Otsu's

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

method maximizes the between-class variance of pixel intensities, effectively separating the foreground from the background.

(e) ***Unet with ResUnet*** - Unet with ResUnet combines the U-Net architecture with residual connections (ResNet). It is a deep learning-based segmentation model that has shown excellent performance in various segmentation tasks. Unet with ResUnet consists of encoder-decoder architecture with skip connections, allowing for precise localization and segmentation.

Step 4: Selection of best Segmentation technique

In this step, the segmentation method that best balances performance, computational efficiency, and ease of implementation for LULC aerial images is selected.

Unet with ResUnet model as a backbone performs best and hence has been selected.

Step 5: Performance evaluation Matrices

The quantitative performance of different DL-based segmentation algorithms on LULC datasets is demonstrated by a few of the metrics that are most commonly used to assess the performance of segmentation models. The efficacy on LULC segmentation has been evaluated using the following metrics: Accuracy, Jaccard Similarity Coefficient (JSC), Dice Similarity Coefficient (DSC), Intersection Over Union (IoU), and Mean Intersection Over Union (mIoU).

Equations (1), (2), (3), (4), and (5) demonstrate how to calculate the accuracy, Jaccard, DSC, IoU, and mIoU. True negative, true positive, false negative, and false positive are represented, respectively, by the symbols TN, TP, FN, and FP.

$$\text{Accuracy} = 1+\left(\frac{TP+TN}{TP+FN+TN+FP}\right) \quad (1)$$

$$\text{Jaccard} = \frac{TP}{TP+FP+FN} \quad (2)$$

$$\text{DSC} = \frac{2\,TP}{2\,TP+FP+FN} \quad (3)$$

$$\text{IoU} = \frac{TP}{TP+FP+FN} \quad (4)$$

$$m\text{IoU} = \frac{\sum_{i=0}^{n} IoUi}{N} \quad (5)$$

7. IMPLEMENTATION & RESULTS

7.1 Dataset description

Ten distinct classifications can be seen in the high-resolution land use and land cover photos included in the dataset. Every picture is 64 by 64 pixels in size, with a ground

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

sampling distance of 10 meters. For the segmentation job in this study, aerial photos of ten different classes are provided. The RGB photos in this collection are taken from the Sentinel Dataset. Ten different classifications make up this dataset, which satisfies the demand for a more extensive range of land-use and land-cover characteristics from different provinces and regions. As seen in Fig 2, the Sentinel-2 satellite was the source of all of these photos.

The following datasets fields are present in the directories:

1) Forests
2) Highway
3) Herbaceous Vegetation
4) Agricultural land
5) Industrial
6) Sea & Lakes
7) Permanent Crop
8) River
9) Residential
10) Pasture

Fig 2: Study- images captured by satellites

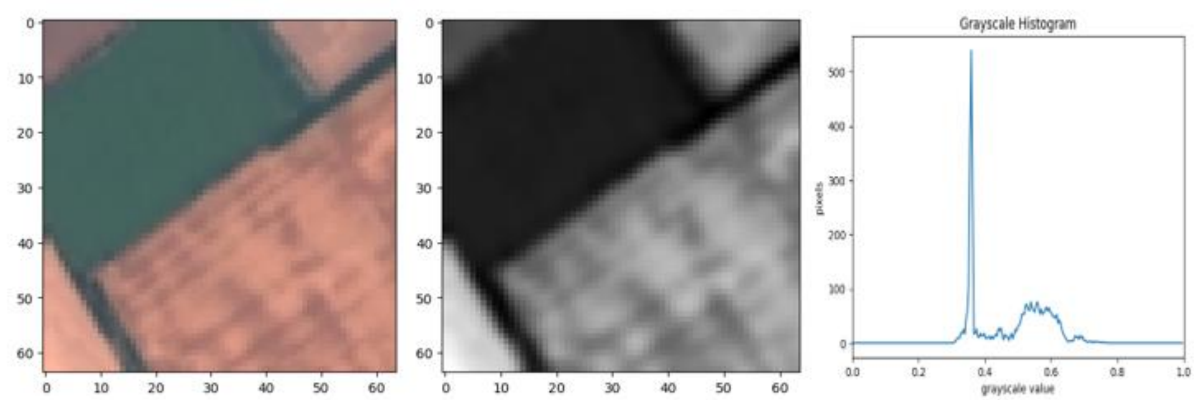
The 64*64 pixel images in the collection have a resolution of 10 centimeters. For every kind of land use, five distinct image segmentation methods have been used. Each land-use class's model employed a total of 2520 and 1080 photographs, with training and testing utilizing a random ratio of 70% (280 photos) to 30% (120 images).

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

7.2 Technique Used

7.2.1 Threshold based Segmentation

This method divides an image into smaller sections based on variations in the image's grey scale values. By choosing a threshold value, it may also be used to separate foreground objects from the background. One can convert a grey level image to a binary image. The binary image should contain all pertinent information about the shape and location of the things of interest. Getting a binary image has the benefit of simplifying the recognition process and lowering data complexity. It is challenging to identify notable peaks and troughs in the photograph (Salwa Khalid Abdulateef 2021). Its shortcomings include noise sensitivity, difficulty setting thresholds, and lack of consideration for spatial detail, which could result in segments that are not contiguous (Abubakar 2013). A further drawback is to the worry of computational complexity, which escalates in direct relation to image size (Garg 2016). An example of threshold-based segmentation is shown in Figure 3.



(a) (b) (c)

Fig 3: Threshold-based segmentation (a) original image, (b) the resulting image after segmentation (c) grayscale histogram

7.2.2 Edge based Segmentation

In this class, segmentation is carried out by locating the image's edges on its gradient in order to determine the objects' boundaries. This method uses edges as a criterion for object identification. These objects, which comprise the optimal edge detector, the second derivative operator, and the first order derivative/gradient operator, are typically identified by it (Salwa Khalid Abdulateef 2021).

1) Gradient operators, which work with first-order derivatives, react to changes in intensity level discontinuities. It has two edges: a positive leading edge and a negative following edge. Prewitt, Roberts, and Sobel operators can identify edges by determining the magnitude of the first derivative. (B. K. Shah, 2020), (Chakraborty, 2020).

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

2) When the lighter side is negative and the darker side is positive, use the second derivative operator. It is very susceptible to picture noise. However, it is great for gathering some secondary data, such the Laplacian operator and the Difference of Gaussian (DoG), which finds edges by looking for zero-crossings (S. E. R. T. Eser, 2019).

3) As a perceptive edge detector, the optimal edge detector may generate noise-resistant, continuous edges that are only one pixel thick. It can also identify edges that are weak and strong. (B. K. Shah, 2020), (D. Sangeetha, 2019).

There are other reasons why the edge-based section is poor. The first reason is that the method is unable to produce the objects' boundaries. The rationale is that it permits incomplete and jumbled borders between the background and the region of interest. The edges created by this method were discontinuous in the majority of cases. Its sensitivity to noise is the second factor. Consequently, the edge often failed to accomplish proper segmentation for noisy pictures. Its sensitivity to the expansion of edges between significant regions is the third component. Images with intricate geometric shapes and high spatial resolution present these challenges. As a result, there is little definition of the boundaries between the zones, resulting in either over or under segmentation. (Salwa Khalid Abdulateef, 2021). An example of edge-based segment is shown in Fig 4.

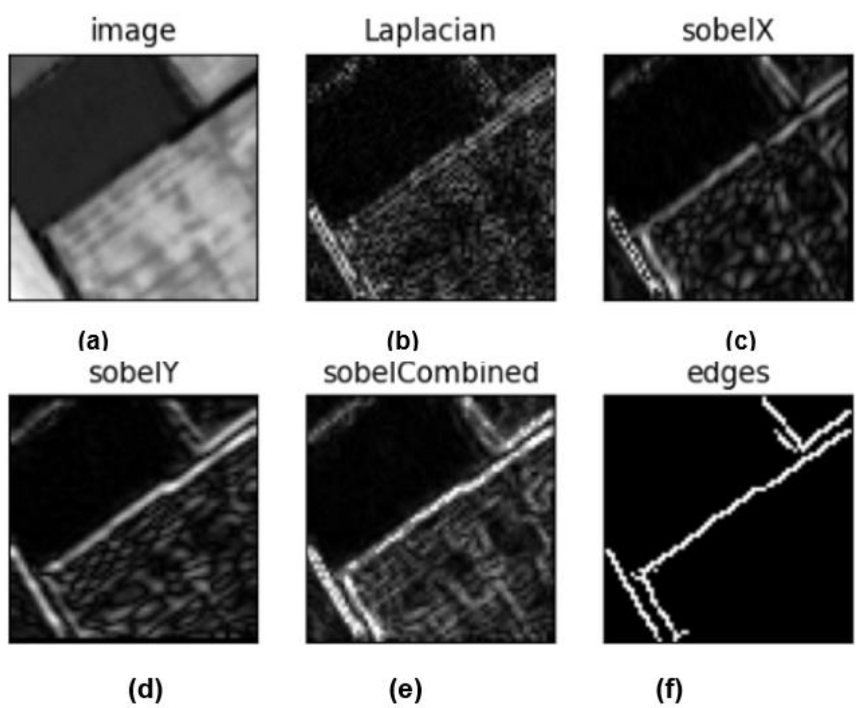


Fig 4: Edge-based segmentation (a) original image, (b) resulting image after segmentation using Laplacian Operator, (c) Sobel Operator (Sobel X), (d) Sobel Operator(Sobel Y), (e) Sobel Combined, (f) resulting image after segmentation using optimal edge detector

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

7.2.3 K-means Clustering based Segmentation

When categorizing picture pixels into k number of clusters—where k is a whole positive number—the K-means technique is employed. The K-means technique was developed by Macqueen (J. Xu 2019) and has been applied to picture segmentation in a variety of fields, including medical (K. Shrivastava 2014), caloric calculation (T. Ege 2019), and S. Turmchokkasam (2018). Based on several similarity features, including pixel intensity, colour, and distance, this categorization is carried out (H. Xiao 2021). It's crucial to note that one benefit of clustering is that it doesn't require any prior knowledge of the data's distribution. Researchers have been drawn to use k-means for picture segmentation due to its efficiency, simplicity, ease of implementation, and speed at which large numbers of data points may be clustered. This approach has a number of problems, including a significant dependence on the initial conditions, inconsistent outcomes when repeated, and an incorrect definition of the number k that is needed. The k-mean method is not without its drawbacks, though. is noise-sensitive, to start. Second, the cluster's choice numbers are restricted. Thirdly, distinct beginning centroids yield distinct results and require more computational effort, hence lengthening the computational duration (Salwa Khalid Abdulateef 2021). An example of segmentation based on K-means clustering is shown in Fig 5.

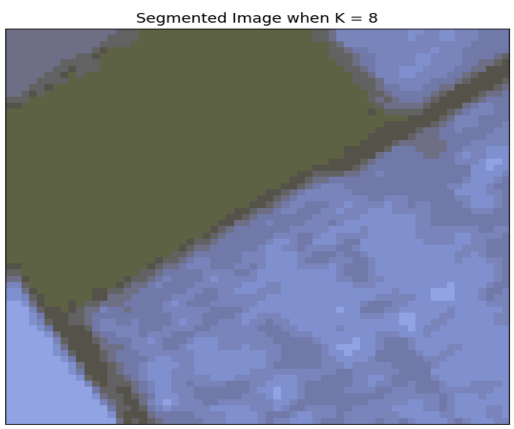


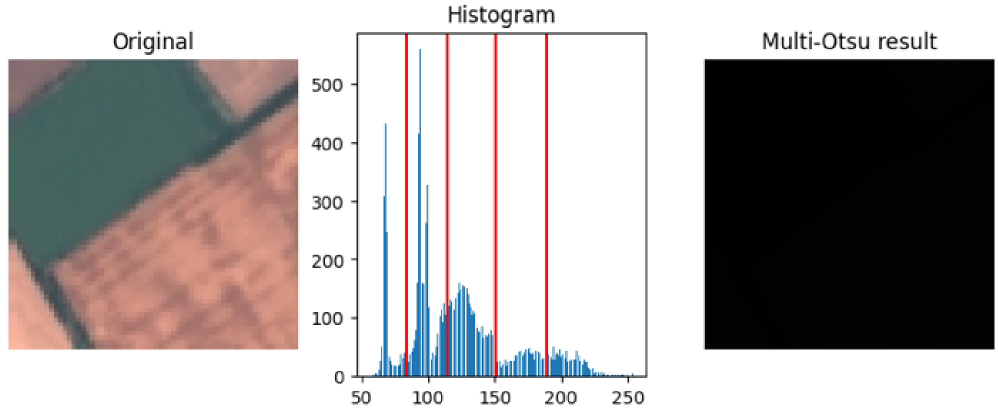
Fig 5: Illustrates the resulting image after K-means clustering based segmentation

7.2.4 Otsu's Segmentation

Otsu's approach finds the ideal threshold value that divides the foreground and background regions with maximal inter-class variance using the image's grey scale histogram. Otsu's technique works well for easy image thresholding jobs. The simplicity and quickness of Otsu's approach are just two of its numerous advantages. The best threshold value to distinguish the foreground from background portions of the processed image can be automatically determined, negating the need for prior knowledge of the image. It also functions well with bimodal histograms, which are typical in a wide range of applications. Otsu's approach can only calculate one threshold value, hence it performs poorly with photos that have histograms with more than two peaks. Its assumption that

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

the variances of the foreground and background regions are equal may not always hold true, resulting in subpar segmentation outcomes. It gives erroneous results for photos with unequal illumination and lighting. It is not noise-resistant, which may cause erroneous thresholding outcomes. An example of Otsu's segmentation is shown in Fig 6.



(a) (b) (c)

Fig 6: Otsu's segmentation (a) original image, (b) Histogram, (c) the resulting image after applying Otsu's segmentation

7.2.5 Unet with ResUnet as a Backbone model

In an effort to enhance performance on image segmentation tasks, ResUNet is a convolutional neural network architecture that blends residual connections with the U-Net architecture. It combines the advantages of residual connections for better feature learning and propagation with the strengths of U-Net's design for accurate segmentation.

- ***Input Image***

This is the initial stage where the image to be segmented is fed into the network. It's typically represented as a 2D or 3D array of pixel values.

- ***Encoder (Contracting Path)***

In this stage, the input image passes through a series of convolutional layers, often with pooling layers interspersed between them. Each convolutional layer captures features from the image, starting from low-level details to high-level semantic information. Down-sampling operations like max-pooling reduce the spatial dimensions of the feature maps while increasing the number of channels, facilitating hierarchical feature extraction.

- ***Bridge***

The bridge connects the encoder to the decoder. It usually consists of additional convolutional layers without down-sampling. The purpose of the bridge is to bridge the semantic gap between the encoder and decoder parts of the network, allowing for the capture of more context-rich features.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

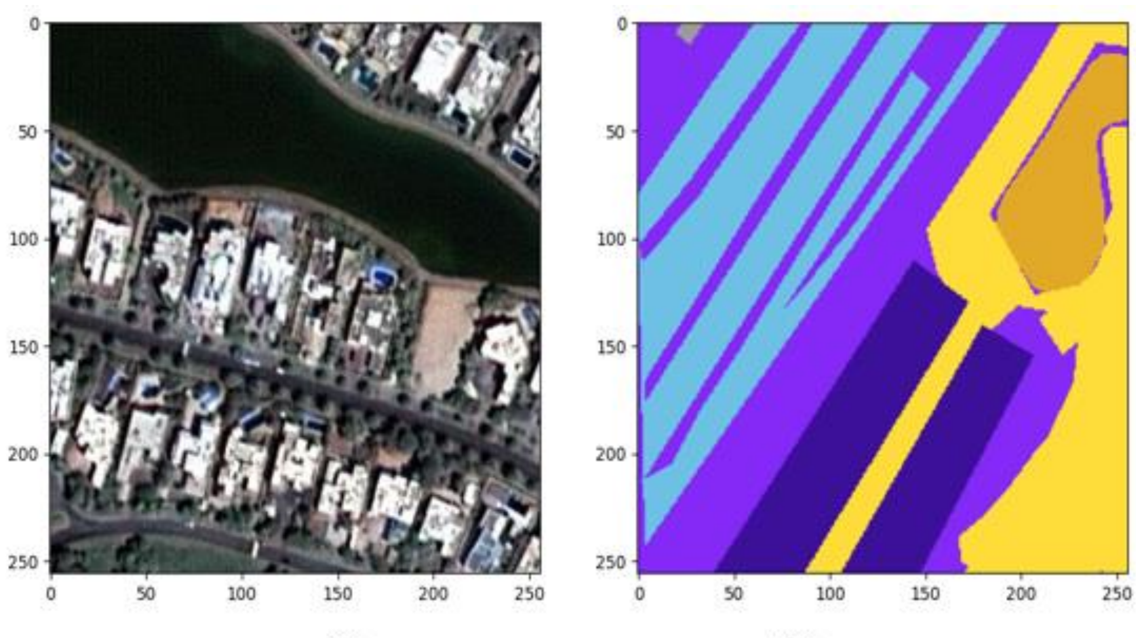
- ***Decoder (Expanding Path)***

The decoder receives the output from the bridge and performs the opposite operation of the encoder. It involves up-sampling the feature maps back to the original input size while reducing the number of channels. Each up-sampling block is typically composed of up-sampling operations like transposed convolutions or bilinear interpolation, followed by convolutional layers. Skip connections from the encoder are concatenated with the corresponding decoder feature maps to preserve fine-grained details and spatial information lost during down-sampling. These skip connections are crucial for precise localization and segmentation, as they allow the network to directly access low-level features from the encoder while combining them with high-level features from the decoder.

Application of Unet with ResUnet as a Backbone model on images: - In Fig 7(a) aerial image of residential land use is provided, Fig 7(b) the resulting image after applying image segmentation to residential land use is provided. In Fig 8(a) aerial image of industrial land use is provided, Fig 8(b) the resulting image after applying image segmentation to industrial land use is provided. In Fig 9(a) aerial image of highway is provided, Fig 9(b) the resulting image after applying image segmentation to highway is provided. In

Fig 10(a) aerial image of river is provided,

Fig 10(b) the resulting image after applying image segmentation to river is provided.



(a) (b)

Fig 7: Unet with ResUnet as a Backbone model for segmentation (a) Original image of residential land use, (b) the resulting image after segmentation

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

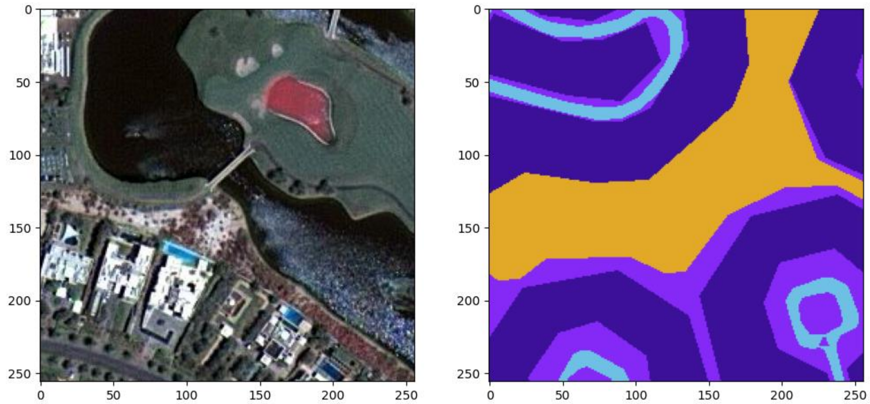
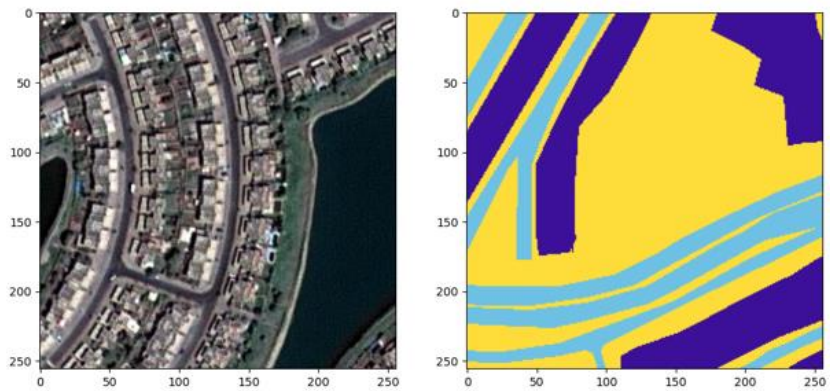
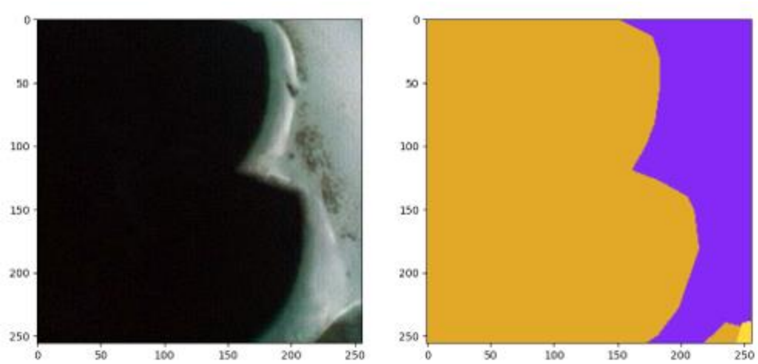


Fig 8: Unet with ResUnet as a Backbone model for segmentation (a) Original image of industrial land use, (b) the resulting image after segmentation



(a) (b)

Fig 9: Unet with ResUnet as a Backbone model for segmentation (a) Original image of highway, (b) the resulting image after segmentation



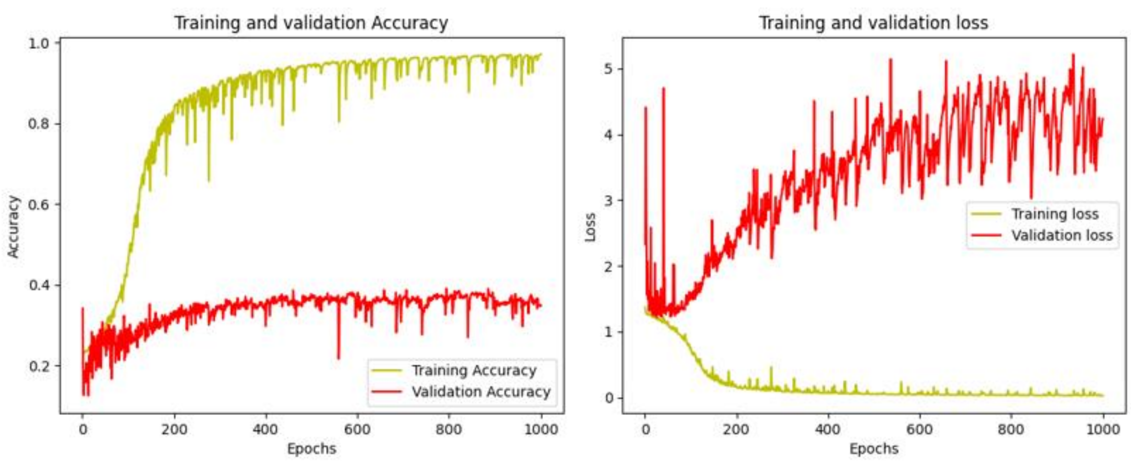
(a) (b)

Fig 10: Unet with ResUnet as a Backbone model for segmentation (a) Original image of river, (b) the resulting image after segmentation

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

The model's ability to predict the right labels on the training dataset is gauged by the training accuracy metric. It is the ratio of the total number of examples in the training set to the number of successfully predicted occurrences. The model's ability to predict the right labels on a different validation dataset that it wasn't exposed to during training is measured by validation accuracy. Similar to training accuracy, but it assesses how well the model generalises. A graph illustrating training accuracy and validation accuracy is shown in Fig. 11(a).

The model's prediction error on the training dataset is measured by training loss. By computing the difference between the expected and actual output for each training instance, it measures the model's performance. The model's prediction error on the validation dataset is measured by the validation loss. It serves as a gauge for how well the model should function with fresh, untested data. The training loss and validation loss are shown in Fig. 11(b).



(a) **(b)**

Fig 11: Graphs showing Accuracy and Loss of Training and Validation Dataset (1000 epochs): (a) Training and Validation Accuracy, (b) Training and Validation Loss

We usually utilise a different dataset known as the test dataset to assess a trained machine learning model. A single image from the test dataset is referred to as the "testing image" in Figure 12(a). These pictures are kept apart and aren't included in the training phase to assess how well the model performs with unknown data. The trained model uses each test image as an input to make predictions.

The label for testing The ground truth, accurate classification, or regression value linked to a test image is shown in Fig. 12(b). It depicts the real class or group that the tested image is a part of the output that the trained model produces after processing a test image is shown in Fig. 12(c). Usually, the prediction is a probability distribution over classes or a class label.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

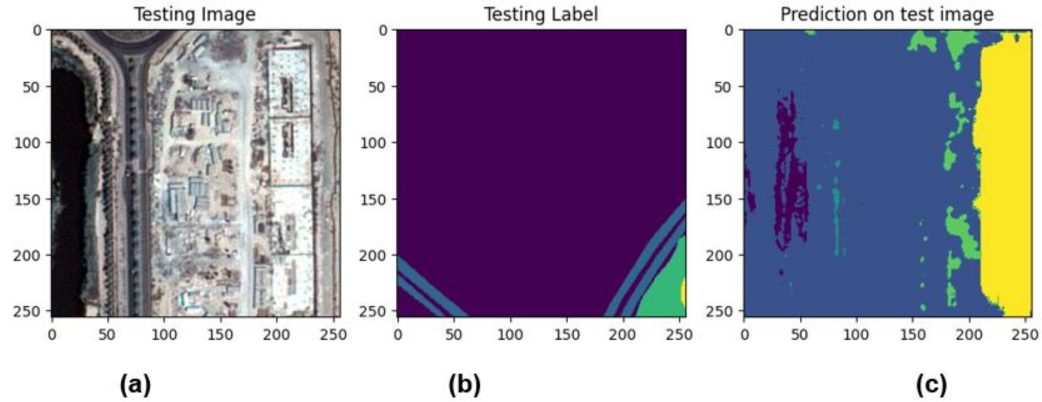


Fig 12 : Results after the Dataset is labelled and ResUnet is applied: (a) Testing Image, (b) Testing label, (c) Prediction on tested image

The two most important indicators for assessing a machine learning model's performance during training are training and validation loss; represented in **Error! Reference source not found.**(a). Each iteration (or epoch) of the training process is measured by the training loss, which indicates how well the model is performing on the training data. It measures the discrepancy between the model's anticipated outputs and the training data's actual target values. The model gets more and more adept at making predictions on the training data as the purpose of training is to minimise the training loss. Training IoU measures the Intersection over Union of predicted objects or segments compared to the ground truth annotations during the training process; represented in **Error! Reference source not found.**(b). Validation IoU measures the Intersection over Union of predicted objects or segments compared to the ground truth annotations on a separate validation dataset. IoU (Intersection over Union) is a common evaluation metric used in tasks like object detection and semantic segmentation to measure the overlap between predicted bounding boxes or segmentation masks and their ground truth counterparts.

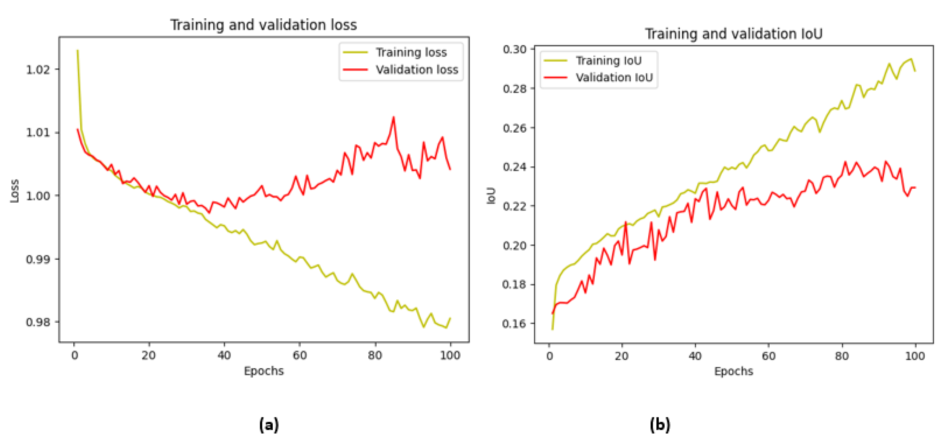


Fig 13: Graphs showing Loss and Intersection over Union for Training and Validation Dataset (100 epochs): (a) Training and Validation Loss, (b) Training and Validation (IoU) Intersection over Union

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

The main result of the ResUNet backbone model-trained U-Net would be semantic segmentation masks (Fig. 14). These masks show the input image's pixel-by-pixel predictions of the land use and land cover classifications. A particular land use or land cover class, such as cities, forests, water bodies, agricultural fields, etc., is represented by each pixel in the output masks. The input image is efficiently segmented into multiple regions corresponding to different land use or land cover categories as the model learns to assign a class label to each pixel.

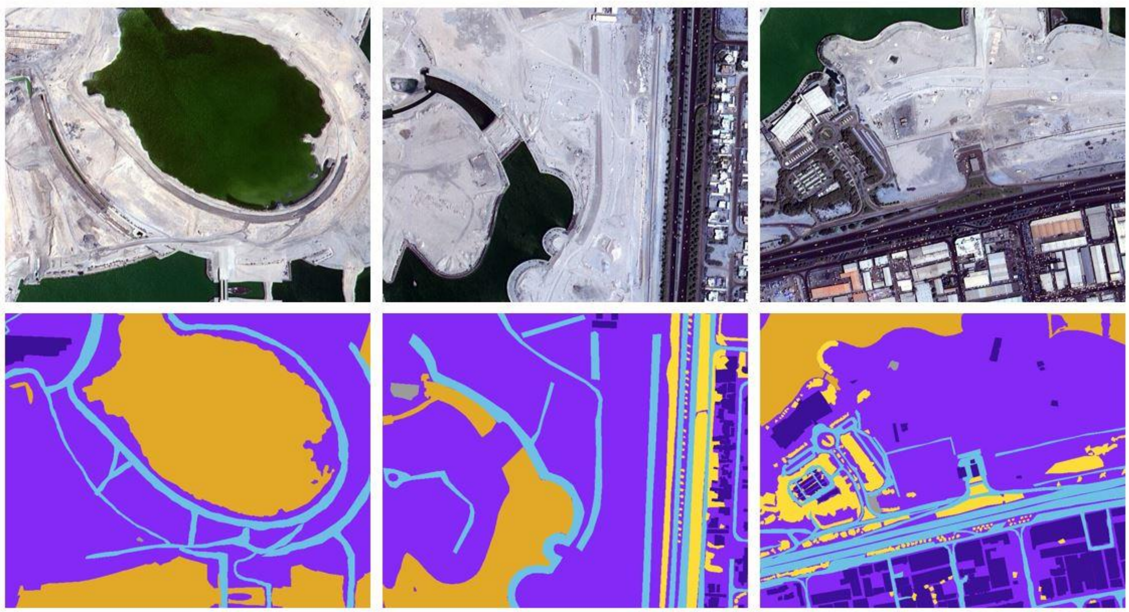


Fig 14: Final Output after Unet with ResUnet as a backbone model is trained, tested and validated on Land use land cover dataset

8. CONCLUSION AND FUTURE SCOPES

This work demonstrates the utilization of remote sensing pictures from many sources, particularly radar and hyper spectral datasets, to segment land use and land cover using image segmentation techniques. The authors also presented a validation procedure for assessing the effectiveness of the artificial neural network. Cross-validation tests using all methodologies on several datasets were statistically examined, and the results showed that the projected artificial neural networks performed better than the other models in terms of overall and per-class accuracy. When segmentation is not used in aerial image analysis, the classification of land cover may not be correct. Mistaken boundary demarcation raises the risk of misclassifying different types of land cover. Aerial imagery covers large areas; therefore it can be difficult to find specific objects or areas of interest without segmentation. It could be more challenging to carry out targeted analysis or monitoring due to this lack of localization. The results indicate that Unet, using ResUnet as the backbone model, performed best for image segmentation of land use and cover, with an accuracy of at least 98.97%, Jaccard coefficient of 97.07%, DSC of 98.49%, IOU of 89.80%, and MIOU of 93.76%. Ultimately, the study demonstrated that artificial neural networks are incredibly powerful tools for segmenting low-latency low-level cinema images. This is because all of the photos under investigation yielded positive results, despite their varied origins and distinctive characteristics.

Furthermore, the article endeavored to illustrate the necessity of a cohesive validation methodology, which might be employed to evaluate the quality of novel proposals for the remote sensing industry. Future research along this line of inquiry necessitates a more thorough examination of the network's design and characteristics. Furthermore, studies on the application of sophisticated post-processing techniques to improve output quality may be conducted. Lastly, a more admirable long-term objective may be to alter the system to enable it to handle data that has been combined from numerous sensors and sources.

Author Contributions

Smita Sunil Burrewar: Conceptualization, methodology, analysis, writing-original draft, **Mazharul Haque:** Review and editing, **Tanwir Uddin Haider:** Conceptualization, methodology. All authors have read and agreed to their specific contribution.

Funding

No funding has been raised for this research.

Data Availability

The authors have no permission to share the data.

Declarations

Competing Interests The authors declare no competing interests.

Ethics Approval Not Applicable.

Consent to participate All authors reviewed and approved the final manuscript.

Consent to Publish All authors approved this for publication.

References

1) Abdusalomov, A., Mukhiddinov, M., Djuraev, O., Khamdamov, U., & Whangbo, T. (2020). Automatic Salient Object Extraction Based on Locally Adaptive Thresholding to Generate Tactile Graphics. Applied Science, 33-50.

2) Abubakar, F. M. (2013). Study of image segmentation using thresholding technique on a noisy image. International Journal of Science and Research (IJSR), 49-51.

3) B. K. Shah, V. K. (2020). Evaluation and Comparative Study of Edge Detection Techniques. IOSR Journal of Computer Engineering, 6-15.

4) Bengana, N., & Heikkilä, J. (2021). Improving land cover segmentation across satellites using domain adaptation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. , 1399–1410.

5) C. Zhu, C. U. (2019). Improved logistic regression model for diabetes prediction by integrating PCA and K-means techniques. Informatics in Medicine Unlocked, 100-179.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

6) Chakraborty, S. (2020). An advanced approach to detect edges of digital images for image segmentation. In Applications of Advanced Machine Intelligence in Computer Vision and Object Recognition: Emerging Research and Opportunities, 90-118.

7) Chhor, G., Aramburu, C., & Bougdal-Lambert, I. (2017). Satellite Image Segmentation for Building Detection using U-net. Computer Science Semantic Sch., 114–120.

8) D. Sangeetha, &. P. (2019). FPGA implementation of cost-effective robust Canny edge detection algorithm. Journal of Real-Time Image Processing, 557-570.

9) Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., et al. (2018). DeepGlobe 2018: A Challenge to Parse the Earth through Satellite Images. CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), (pp. 18-22). Salt Lake City, UT, USA.

10) Garg, A. (2016). A Review on Image Segmentation Techniques. International Journal of Recent Research Aspects (IJRRA), 53-55.

11) H. Xiao, L. H. (2021). Texture synthesis: a novel method for generating digital models with heterogeneous diversity of rock materials and its with heterogeneous diversity of rock materials and its CGM verification. Computers and Geotechnics, 56-78.

12) Huang, G., Liu, Z., & Weinberger, K. (2017). Densely connected convolutional networks. IEEE Conference on Computing and Vision Pattern Recognition., (pp. 2261-2269). Honolulu, HI, USA.

13) J. Xu, &. K. (2019). Power k-means clustering. International Conference on Machine Learning, 6921-6931.

14) K. Shrivastava, N. G. (2014). Medical image segmentation using modified k means clustering. International Journal of Computer Applications, 12-16.

15) Karwowska, K., & Wierzbicki, D. (2022). Improving Spatial Resolution of Satellite Imagery Using Generative Adversarial Networks and Window Functions. Remote Sensing, 62-85.

16) Kuo, T., Tseng, K., Yan, J., Liu, Y., & Wang, Y. (2018). Deep aggregation net for land cover classification. IEEE/CVF Conference on Computing and Vision Pattern Recognition Workshops 2018, (pp. 247-252). Salt Lake City, UT, USA.

17) Kutlimuratov, A., Abdusalomov, A., & Whangbo, T. (2020). Evolving Hierarchical and Tag Information via the Deeply Enhanced Weighted Non-Negative Matrix Factorization of Rating Predictions. Symmetry.

18) Lee, S., Park, S., Son, S., Han, J., Kim, S., & Kim, J. (2019). Land cover segmentation of aerial imagery using SegNet. Earth Resource, Environment, Remote Sensing/ GIS Application., 313-318.

19) S. E. R. T. Eser, &. A. (2019). A new edge detection approach via neutrosophy based on maximum norm entropy. Expert Systems with Applications, 499-511.

20) S. Turmchokkasam, &. K. (2018). The design and implementation of an ingredient-based food calorie estimation system using nutrition knowledge and fusion of brightness and heat information. IEEE Access, 46863-46876.

21) Salwa Khalid Abdulateef, M. D. (2021). A Comprehensive Review of Image Segmentation techniques. Iraqi Journal for Electrical and Electronic Engineering, 166-175.

22) Sertel, E., Ekim, B., Osgouei, P., & Kabadayi, M. (2022). Land Use and Land Cover Mapping Using Deep Learning Based Segmentation Approaches and VHR Worldview-3 Images. Remote Sensing, 45-53.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 09:2024
DOI: 10.5281/zenodo.13744019

23) Sevak, J., Kapadia, A., Chavda, J., Shah, A., & Rahevar, M. (2017). Survey on semantic image segmentation techniques. International Conference on Intelligent Sustainable Systems (ICISS), (pp. 306-313). Palladam, India.

24) Sharifzadeh, S., Tata, J., Sharifzadeh, H., & Tan, B. (2019). Farm Area Segmentation in Satellite Images Using DeepLabv3+ Neural Networks. In Data Management Technologies and Applications., 12-55.

25) Smita.S. Burrewar, M. H. (2024). Convolutional Neural Network Methods for Detecting Land-use Changes. International Journal of Intelligent Systems & Applications in Engineering, 573-590.

26) T. Ege, Y. A. (2019). Image-based estimation of real food size for accurate food calorie estimation. 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), 274-279.

27) Tian, C., Li, C., & Shi, J. (2018). Dense fusion classmate network for land cover classification. IEEE/CVF Conference on Computing and Vision Pattern Recognition Workshops 2018 (pp. 18–22). Salt Lake City: IEEE.

28) Ulmas, P., & Liiv, I. (2020). Segmentation of satellite imagery using U-net models for land cover classification. 22-28.

29) Wafa, R., Khan, M., Malik, F., Abdusalomov, A., Cho, Y., & Odarchenko, R. (2022). The Impact of Agile Methodology on Project Success, with a Moderating Role of Person's Job Fit in the IT Industry of Pakistan. Applied Science., 98-106.