Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 11, No. 3, September 2023, pp. 630~642
ISSN: 2089-3272, DOI: 10.52549/ijeei.v11i3.4662

Detecting Urban Road Changes using Segmentation and Vector Analysis

M. Sobhana[1], Gudapati Satya Dinesh Kumar[2], Yarramreddy Tejaswi[3], Pavithra Pakkiru[4]
[1]Faculty at Department of Computer Science and Engineering, V R Siddhartha Engineering College, India
[2] Department of Computer Science and Engineering, V R Siddhartha Engineering College, India

Article Info

Article history:

Received Apr 6, 2023
Revised Jul 7, 2023
Accepted Aug 15, 2023

Keywords:

Change Detection
Co-registration
Semantic segmentation
U-Net
Vector analysis

ABSTRACT

The rapid growth of urbanization is driving increased road infrastructure development. Detecting and monitoring changes in urban road areas is challenging for city planners. This research proposes using semantic segmentation and vector analysis on high-resolution images to identify road network changes. The U-Net model performs semantic segmentation, pre-trained on a Massachusetts road dataset, predicting labels for a specific area with temporal data and co-registration to reduce distortions. Predicted labels are converted to shapefiles for vector analysis. Satellite images from Google Earth archives demonstrate the change detection process. The outcome of this predictive phase was the transformation of projected labels into shapefiles, thereby facilitating vector analysis to pinpoint and characterize alterations.

Corresponding Author:

Dr. M. Sobhana
Department of Computer Science and Engineering,
Velagapudi Ramakrishna Siddhartha Engineering College,
Email: sobhana@vrsiddhartha.ac.in

1. INTRODUCTION

Many urban applications, including, urban planning, traffic management and road monitoring, rely heavily on remote sensing for road extraction. The occlusions, noise, and intricacy of the surroundings in unprocessed remote-sensing images make it difficult to extract the roads from high-resolution remote-sensing images.[1] Temporal data from two different years of the Massachusetts area were used to carry out the change detection in this study.

The two-year dataset does not have the same reference system, which makes working with temporal data problematic because there may be minor sub-pixel geometry displacements across images due to source, timing, and perspective variations throughout the image acquisition process. Even across images taken with the same sensor, displacements might happen.[2] Thus, to overcome this challenge, this proposed study makes use of the image co-registration technique which is the process of geometrically aligning two or more images and allows for the integration or fusion of matching pixels that represent the same objects. Simply said, the co-registration process rectifies the distortion or deformation of an image concerning a reference image. The reference image has been orthorectified and adjusted for geometry. It indicates that it was georeferenced using ground-based control points and geometry-corrected using knowledge of the terrestrial terrain [3].

An image can be divided into segments by an image processing method called segmentation that enables us to distinguish between objects and textures in it[4]. They are several deep-learning models available for Semantic Segmentation. U-Net which is a widely popular deep learning model is used in this research. This study uses the U-NET architecture with RESNET-50 as a backbone that has been pre-trained on the ImageNet database for image segmentation and to extract road networks from satellite images. The output generated by the U-Net model is raster data which is in a compressed format that can result in data loss. To avoid data loss, raster data need to be converted into vector data so that image data remainssecured. It also improves image compatibility, reduces the file size, increases the image portability, and enhances the image appearance.[5]

An approach for determining how a certain area has changed over the course of two or more time periods is called Geographic Information System (GIS) change detection [6]. Understanding changes in ice sheets, land use, and forest cover is made easier by change detection. "Change detection" is the process of comparing changes between aerial images of the same region taken at distinct periods in time. As a result, Quantum Geographic Information System (QGIS), an open-source software program used mostly for GIS applications, is employed to carry out the change detection.

M. Sahu and A. Ohri [7] proposed a simple yet successful approach for extracting and differentiating individual buildings from densely crowded locales using medium-resolution unmanned aerial vehicle images. The obtained output is put into a post-processing pipeline that has been provided in order to separate related binary blobs of buildings and translate it into a Geographic Information System layer for extra analysis and also to create three-dimensional structures. Jonas Botelho et al. [8] made research that sought to utilize artificial intelligence to automatically identify unofficial roadways in the Brazilian Amazon. In this study, the Azure Planetary Computer Platform’s Sentinel-2 footage from 2020 was used to train and build a modified U-shaped encoder-decoder network algorithm to recognize rural roads in the Brazilian Amazon.

Volodymyr Hnatushenko et al. [9] presented this work, which addresses the problem of co-registering optical satellite images with Synthetic Aperture Radar (SAR). Mohamed Tahounl et al. [10] presented research in which invariant local characteristics were used to offer a coregistration method for satellite images. During the keypoint extraction, descriptor creation, and matching operations, many corners and scale-based feature detectors have been put to the test. Pairwise matching of the input images is obtained, and then a comprehensive registration process that includes implementing bundle correction and warping of an image before composing the version that is registered is carried out.

André Stumpf et al. [11] presented a method to adjust for systematic offsets and striping artefacts. The approach uses dense sub-pixel offset measurements and strong statistics. The effectiveness of the processing chain is assessed experimentally at locations with various environmental circumstances and compared to previously suggested procedures. Chen, Ziyi, et al. [12] proposed extracting roadways from high-resolution optical remote sensing images using a rehabilitation bias U-Net. It is split into two parts: encoding and decoding. Liqin Huang et.al [13] created a road central point estimation technique that was surface-based and can estimate highway central points in each area. By using this method, you may avoid the massive concentration of pixel-by-pixel labels required for feature-based model training.

Abdollahi, Abolfazl, et al. [14] carried out an in-depth analysis of deep learning algorithms applied to traditional remote sensing criteria for highway extraction. Ostadabbas, Weippert, and Behr [15] used software such as QField and QGIS along with the Postgres database to determine the quality and impact on the region before and after rehabilitation. Martino Terrone et al. [16] combine maps from the beginning of the nineteenth century and maps from the current century to detect the volumetric changes of landforms made by man in a populous region in Genoa during the last two decades.

Felix Dahle et al. [17] described a method that speeds up the process of change detection and updation in digital maps. Unlike traditional classification algorithms to perform change detection, it uses 2.5D data and transfers the observed changes to a digital map. Andrea Tassi et al. [18] presented their work aimed at developing an Object-Oriented Land Use-Land Cover approach that overcomes the drawbacks of existing approaches like difficulty in combining proper functions and tuning various parameters. This approach includes both object textural analysis and object segmentation.

Daniel Scheffler et al. [19] presented AROSICS a python-based software for the correction and detection of misalignments in pixels in remote sensing datasets. It makes use of the Fourier shift theorem, based on phase correlation, to estimate sub-pixel shifts in the frequency domain. Francesca Bovolo et al. [20] introduced a unique unsupervised Deep Change Vector Analysis to identify changes in VHR images. An automated feature extraction is implemented at each layer which selects features having both high and low prior change information. The deep change vectors are observed with their magnitude as a parameter to identify changed pixels.

Kyungsun Lim et al. [21] proposed a change detection algorithm using CNN which eliminates the task of pre-processing like ortho rectification and classification. CNN with three encoders and decoders was developed that generates change maps from satellite images. The result, which is the average of the output from three neural networks, is a binary change map.

A. Research objective

The research aims to develop a methodology for detecting changes in urban road networks using a combination of semantic segmentation and vector analysis applied to high-resolution satellite images. The primary objective is to enhance the ability of city planners to identify and monitor alterations in road infrastructure resulting from rapid urbanization. This involves utilizing the U-Net model, pre-trained on a specific road dataset, to predict labels for temporal data of a designated urban area. The predicted labels are then transformed into shapefiles for vector analysis, allowing for accurate characterization of changes in the road network. The ultimate goal is to provide an effective tool for informed decision-making in urban planning and development.

B. Literature Review

The review is summaried in Table 1. Le-Anh Tran et al. [34] discussed a novel approach for road lane markings detection using a semantic segmentation network based on U-Net architecture. This method employs the Hough Transform to identify lines in the segmented images and utilizes K-means Clustering to select the most appropriate lane markings. The system's effectiveness is validated through testing in the CARLA simulator for autonomous driving research, demonstrating promising results in lane detection.

Chen et al. [35] explored the relationship between road network structure and ride-sharing accessibility, using network science measures like degree and closeness centrality. Findings reveal that factors such as degree centrality and closeness centrality, along with population and road density, influence Uber accessibility in Atlanta. The research highlights the importance of considering time windows and utilizes OpenStreetMap for reproducible and scalable analysis, contributing to our understanding of urban networks and ride-sharing accessibility.

J. A. Gómez et al. [36] studied that urban growth modeling is vital for urban planning and policy-making. He compared two models, SLEUTH CA and a machine learning framework, to analyze spatiotemporal uncertainty in urban land predictions. The study focuses on Jiaxing and Lishui in China, highlighting the importance of considering uncertainty in urban planning for more effective decision-making.

Table 1. Comparative Analysis

Study	Focus	Methodology	Key Findings
Kyungsun Lim et al. [21] (2018)	Change detection algorithm for high resolution satellite images	Deep learning (CNN)	Achieved accurate change detection using CNN-based approach.
Le-Anh Tran et al. [34] (2019)	Semantic segmentation for urban planning	U-Net model	U-Net effectively segmented urban features, aiding in city planning.
Chen et al. [35] (2020)	Road network analysis	Vector analysis	Combined vector analysis with satellite imagery for comprehensive road network assessment.
J. A. Gómez et al. [36] (2021)	Urban Growth Modeling Uncertainty Analysis	Machine learning	Uncertainty in urban planning for more effective decision-making.
Ostadabbas et al. [15] (2020)	German Urban Development Law and GIS for Effective Planning	Geospatial analysis	Highlighed the use of open-source GIS tools like QGIS and QField for efficient urban planning and legal accounting in the context of German Urban Development Law (StBauFG).

2. RESEARCH METHOD

The architecture of the research consists of data collection followed by pre-processing of the dataset. Later semantic segmentation was performed for road extraction. The raster outputs obtained were converted into vectors and then they were subjected to change detection.

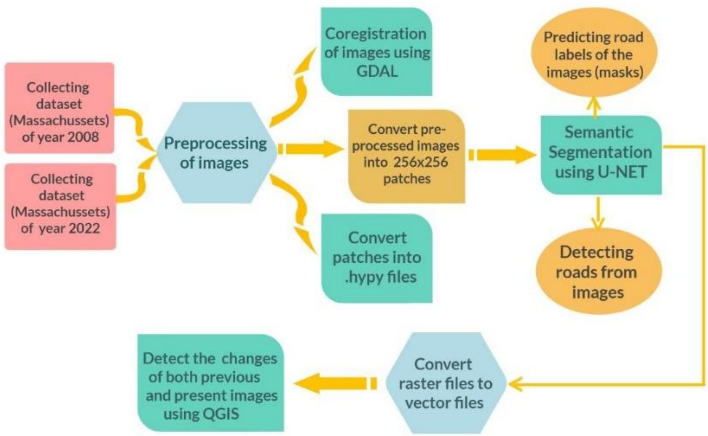

Figure 1. Methodology Diagram

Figure 1 represents the various steps involved and the flow of the steps involved in the research. Dataset collection includes collection temporal data of same area, pre-processing step consists of co-registration, patchify, and file conversion. They are followed by semantic segmentation and and raster to vector conversion of outputs.

2.1. Dataset

The Massachusetts roads dataset is the one that was adopted in the study to train the model. This dataset comes from the original Massachusetts roads dataset created by Volodymyr Mnih. Massachusetts Roads Dataset is a collection of 1171 aerial images of Massachusetts state. Each image has a dimension of 1500 by 1500 pixels and spans 2.25 square kilometers. A training set of 1108 images, a validation set of 14, and a test set of 49 images were randomly selected from the total number of images. More than 2600 square kilometers of urban, suburban, and rural areas are included in the dataset. Almost 110 square kilometers are covered by the test set alone. The target maps were produced using rasterized road centerlines from the OpenStreetMap project. The labels were produced using a 7-pixel line thickness without any smoothing. A resolution of 1-pixel square meter is applied to all images.

To perform change detection, separate satellite images of a specific area are chosen from the Massachusetts area. The images are taken from open source google earth archives. The images are taken from the years 2008 and 2022 to perform change detection.

2.2. Pre-Processing

In this study, two pre-processing approaches were used on satellite images for change detection analysis using the U-Net model. The first pre-processing technique involves co-registration and the second technique is cropping a single large image into smaller patches of size 256x256x3, and transforming the images and masks into HDF5 files.

2.2.1. Co-Registration

In satellite imaging, co-registration refers to aligning multiple images taken by different sensors or at different times in the same area. This is accomplished by identifying common features in the images and adjusting the position and orientation of the images so that these features match up. Co-Registration is useful for applications such as monitoring a region over time, making accurate maps, and discovering data trends. It can be done manually or using automated software tools.

To perform the Co-Registration, a customized Python package called Arosics was employed for calculating shifts and performing the shifting in the target picture. The approach computes the cross-power spectrum of the two pictures using the Fast Fourier Transform (FFT), and from this, the shift may be determined. The sub-pixel shifts are then calculated by fitting a quadratic curve to the cross-correlation peaks surrounding the integer shifts. This is a standard strategy for enhancing shift estimation accuracy.

Figure 2. Satellite Images of the same area in Massachusetts

Figure 2 represents the satellite images of the same area in Massachusetts. The left image was of the year 2008 which has not undergone Co-Registration. And the right image in Figure 2 represents the satellite image from the year 2022. The spatial shifts and pixel shifts should be done before working with temporal data to maintain data integrity. Moreover, Figure 3 represents the same 2018 image of the Massachusset area which has undergone Co-registration using the Arosics python library.

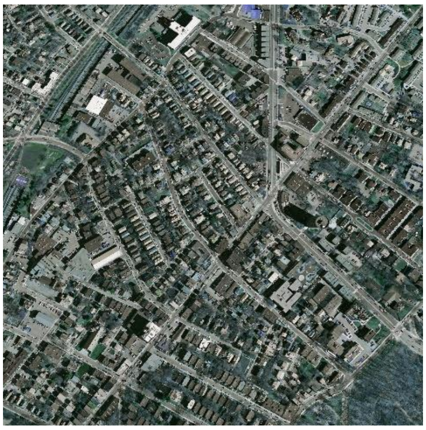

Figure 3. 2008 After Co-Registration

2.2.2. Cropping a single large image into patches of size 256x256x3 for both the images and masks

The U-Net model used in this study requires input images of size 256x256, so the large satellite images were divided into smaller images of this size using the patchify package in Python. The patchify package provides a convenient way to crop the images into smaller patches with a specified size and stride. This was used to divide the images and masks into multiple patches of size 256x256 with a stride of 256 pixels to avoid overlapping while cropping. The patchify package was used to crop the images and masks together to guarantee that the corresponding parts of the images and masks were retained. This was accomplished by feeding both the images and masks into the patchify function, which synchronizedly separated both sets of data into smaller patches.

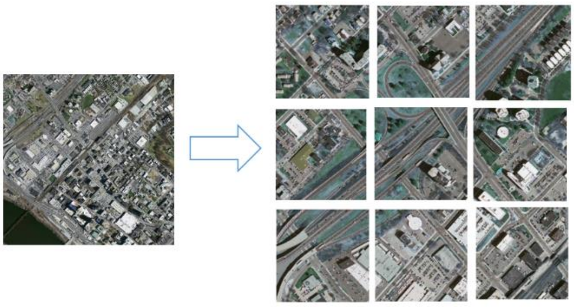

Figure 4. Single large image after being patchfied

2.2.3. Converting the images and masks into HDF5 files:

The HDF5 (Hierarchical Data Format 5) data format is commonly used for processing and storing massive amounts of scientific data, such as satellite images. The input data was converted to HDF5 format using the Python h5py function to get the frames and masks prepared for analysis. For reading and writing data to and from HDF5 files, the h5py package provides a user interface. The input data was organized into datasets and stored in an HDF5 file, with each dataset standing for a single image or mask. This made it simple to retrieve the U-Net model's data for change detection analysis. The HDF5 format was also chosen because it was quick and efficient at storing and retrieving huge volumes of data, which increased the effectiveness and speed of the transformation. The resulting HDF5 files can then be used as input for the U-Net model for semantic segmentation.

2.3. Model Preparation

The U-Net model used in this study is a deep learning model that is specifically designed for image segmentation tasks. The input image size of 256x256x3 (3 represents the number of color channels in the image) was chosen because it provides a good balance between processing speed and accuracy. A larger image size would result in more detailed information being captured, but would also slow down the processing time. On the other hand, a smaller image size would result in faster processing times, but with a loss of detail. The

256x256x3 size allows the model to capture sufficient detail while still operating in an efficient manner. Furthermore, the U-Net model is designed to have a large number of filters, which allows it to effectively extract features from the input images and accurately identify changes in the urban roadways.

The implementation of U-Net utilized is based on the encoder-decoder architecture that features skip connections [22]. Four convolutional blocks comprise the encoder path, with every block containing two 3x3 convolutional layers followed by a max pooling layer. After each convoluted layer comes batch normalization and ReLU activation; this help double up filters in every subsequent block. The decoding block also contains four transposed convolutional blocks. Each one has two 3x3 convolutions supported by concatenation from their associated encoded block through which they get their respective skip connections. Similarly, as in the encoder block, the convolution layers are followed by normalization and ReLU activation. Each block reduces the number of filters by a factor of two. The decoder path produces a 1x1 convolutional layer using a sigmoid.

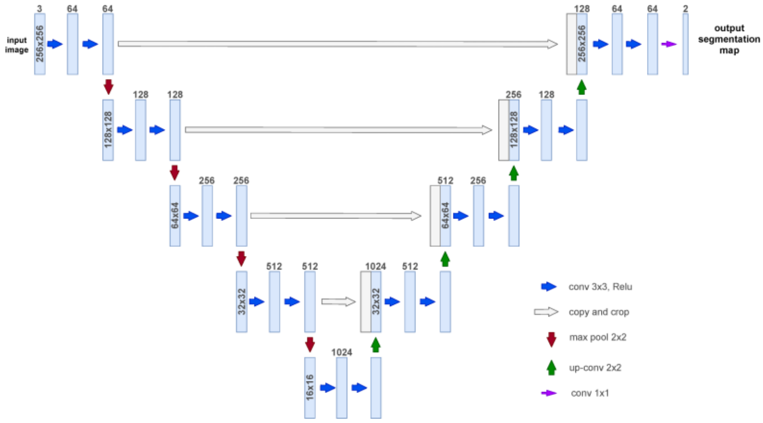

Figure 5. Base U-Net Architecture

The U-Net model works on the basis of convolutions which are a type of linear transformation that is applied to an input image to extract local features. Convolutions can be thought of as sliding a small filter (also known as a kernel or a weight matrix) over an image, multiplying the filter element-wise with the pixels it overlaps, and summing the resulting products to produce a single output value. As mentioned above U-Net has two blocks namely encoder and decoder. The encoder network, which is used to extract high-level features from the input image, is often a pre-trained CNN, among them, ResNet50 has the best performance compared to other backbones [23]. The decoder network is a mirror image of the encoder network that uses up-sampling layers to recover the image's original resolution.

In order to construct the segmentation mask, the U-Net architecture first extracts information from the input image using convolutional layers. The input image and the segmentation mask are labeled as X and Y, respectively. The convolutional layers are denoted by W, and the biases are denoted by b. The activation function is denoted by g.

The convolution operation can be represented using the following equation:

$$Z = X * W + b \tag{1}$$

where * stands for convolution operation, and Z is the output feature map.

Each convolutional layer in U-Net is often followed by a non-linear activation function, such as ReLU, which is denoted by g. The output of the activation function is denoted by A.

$$A = g(Z) \tag{2}$$

where Z is the output of convolution operation from Eq 1.

The U-Net architecture additionally employs pooling layers to reduce the spatial resolution of the feature maps. The most common pooling operation is max pooling, which selects the maximum value in each pooling window. The output of the pooling layer is denoted by P.

$$P = \max\_pooling(A) \tag{3}$$

Where A is the output from the activation function from Eq 2.

The decoder end of the U-Net architecture employs up-sampling layers to improve the feature maps' spatial resolution. The most common up-sampling operation is bilinear up-sampling, which inserts zeros between the input pixels and then applies a low-pass filter to interpolate the missing values. The output of the up-sampling layer is denoted by U.

$$U = bilinear\_upsampling(A) \tag{4}$$

Where A is the output of activation function taken from Eq 2.

Skip connections are also used in the U-Net architecture to combine high-level encoder features with low-level decoder characteristics. The skip connection applies a convolutional layer after concatenating the feature maps from the associated encoder and decoder layers to create the output feature map. In the skip connection, S stands for the output.

$$S = g\big(conv\big(concatenate(A_{encoder}, U_{decoder})\big)\big) \tag{5}$$

where the feature map that results from the corresponding encoder layer is designated as A encoder, and the feature map that results from the corresponding decoder layer is designated as U decoder.

The output feature map of the last decoder layer is then subjected to a convolutional layer to produce the output segmentation mask. A ReLU activation function is used to adjust the convolutional layer's output to a value between 0 and 1. Y_ pred stands for the ReLU activation function's output.

$$Y_{pred} = ReLU(conv(S_{decoder})) \tag{6}$$

where $S_{decoder}$ is the result feature map from the last decoder.

2.4. Model Validation

The images and labels from the Massachusetts dataset are used to train the U-Net model. IoU Score is used as an accuracy metric for training and validation, while soft dice loss is used as a loss metric. Using intersection over union (IoU), mean Average Precision (mAP) is calculated. The degree of overlap between the anticipated and actual bounding boxes is indicated by a value between 0 and 1. The equation depicting the IoU_Score is shown in Equation 1 below.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union(Ground\ Truth\ and\ pred\_mask)} \tag{7}$$

The hyperparameters are chosen in such a way that the model converges as well as possible while adhering to the resource limitations. The model is compiled using the Adam optimizer. The model is trained for 50 epochs, and to avoid overfitting, callbacks such as EarlyStopping, ReduceLROnPlateau, and ModelCheckpoint are used to stop and save the model when it is overfitting. ReduceLROnPlateau's patience level is set to 5 with a factor of 0.8 when the model is not improving. When the val_loss does not improve in subsequent epochs, the patience for EarlyStopping is set to 9. The model is trained with a batch size of 32 and a low learning rate of 0.0001. The model is trained with all the mentioned parameters and the model was able to validate the validation images with a decent validation IoU_Score of 89.4.

To make better use of the predicted labels obtained from road segmentation, which are in raster format, they should be converted into a vector format so that the object of interest, in this case, roads can be selected and overlayed over other predicted labels. Vectorization is performed using the polygonize function available in gdal library of python. Because the roadways have both length and breadth dimensions, polygons were employed to represent the roads as connected polygons.

2.5. Post-processing

The model is made to produce labels for both the year dataset images i.e 2008 and 2022. The labels generated by the U-Net model are of size 256x256x1 which are unpatchified to their original sizes (1500x1500x1). The outputs are in raster format which must be converted into vector format to perform change detection of roads in images.

2.5.1. Vectorization

Vectorization is the process of transforming raster images into vector images. Road networks are frequently well-defined linear features that may be precisely depicted using the mathematical equations that comprise vector graphics.

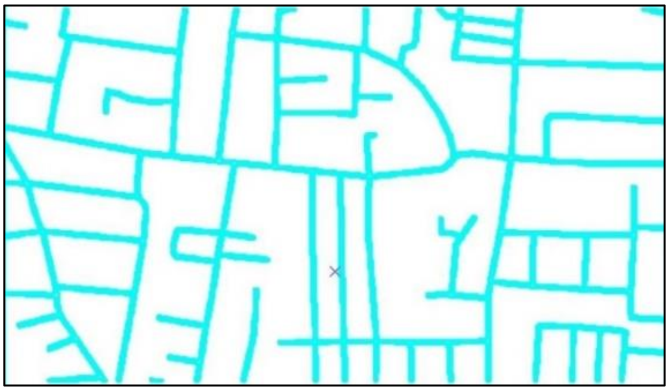
Figure 6. Vectorized mask

To make better use of the predicted labels obtained from road segmentation, which are in raster format, they should be converted into a vector format so that the object of interest, in this case, roads can be selected and overlayed over other predicted labels. Vectorization is performed using the polygonize function available in gdal library of python. Because the roadways have both length and breadth dimensions, polygons were employed to represent the roads as connected polygons.

2.5.2. Change Detection

The vectorization step is applied to both the 2008 and 2022 labels predicted by the U-Net model. The vector files of both labels are added to QGIS in which changes in roads are detected. The open-source QGIS provides various geo-processing tools to perform various set operations on the vector files. The geoprocessing tool namely symmetrical difference is used to identify and eliminate the common features and represent only unsymmetrical elements in both vectors.

2.5.3. Methodology Description: Semantic Segmentation and Vector Analysis

Rationale:

The selection of semantic segmentation and vector analysis methods stems from their complementary nature in addressing the research objective of detecting changes in urban road networks. Semantic segmentation is well-suited for pixel-level classification, enabling the differentiation of road areas from their surroundings. On the other hand, vector analysis facilitates a holistic examination of spatial relationships and quantifies changes in road network geometry over time.

Semantic Segmentation:

Semantic segmentation is implemented using the U-Net model, a convolutional neural network (CNN) architecture widely recognized for its effectiveness in image segmentation tasks. The choice of U-Net is based on its ability to capture intricate spatial features, making it apt for delineating complex road patterns within high-resolution satellite images.

Steps in Semantic Segmentation:

Data Preprocessing: The study employs a dataset of high-resolution satellite images capturing urban areas. These images are preprocessed to normalize lighting conditions, correct distortions, and enhance feature visibility.

Model Training: The U-Net model is pre-trained on a relevant road dataset, such as the Massachusetts road dataset, to learn road-specific features and patterns. Transfer learning is leveraged to fine-tune the model on the study's specific urban area of interest.

Label Prediction: For each temporal instance of the study area, the pre-trained U-Net model predicts pixel-level labels. This entails classifying each pixel as either road or non-road, effectively generating a road map for that specific time.

Vector Analysis:

Vector analysis involves converting the predicted semantic segmentation labels into vector-based shapefiles, which represent road segments as distinct geometric entities. This enables a comprehensive assessment of road network changes, including alterations in road lengths, angles, and intersections.

Steps in Vector Analysis:

Label Conversion: The pixel-wise predicted labels obtained from semantic segmentation are transformed into binary masks, where road pixels are assigned a value of '1' and non-road pixels a value of '0'.

Shapefile Generation: Using the binary masks, shapefiles are generated by vectorizing the road pixels. This process results in individual road segments represented as polygons with associated attributes, such as road type and length.

Temporal Comparison: Shapefiles from different time instances are compared and analyzed spatially. Changes in road segment attributes, such as shape, length, and position, are quantified using geospatial tools and techniques.

3. RESULTS AND DISCUSSION

The semantic segmentation task was performed on the Massachusetts Road Dataset using the U-Net architecture. The model was trained on a subset of the dataset consisting of 1000 images and validated on a separate set of 200 images. The training was carried out for a total of 50 epochs. The performance of the model was evaluated using the intersection over union (IoU) metric and the loss function. The results of the training and validation for both the IoU and loss are shown in Figure 7 and 8, respectively.

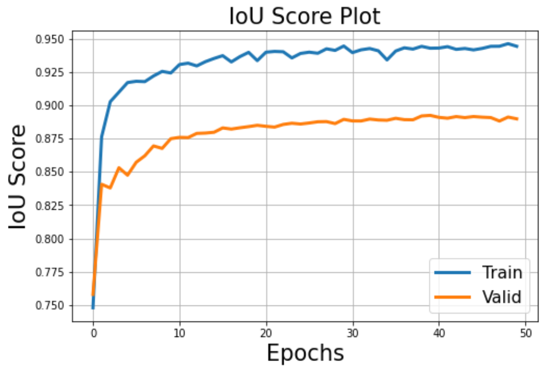

Figure 7. IoU_Score Plot for Training and Validation

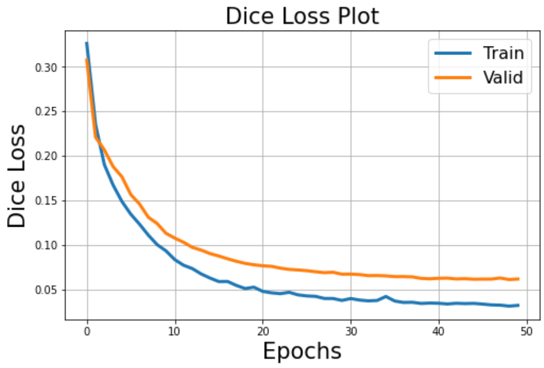

Figure 8. Dice_Loss Plot for Training and Validation

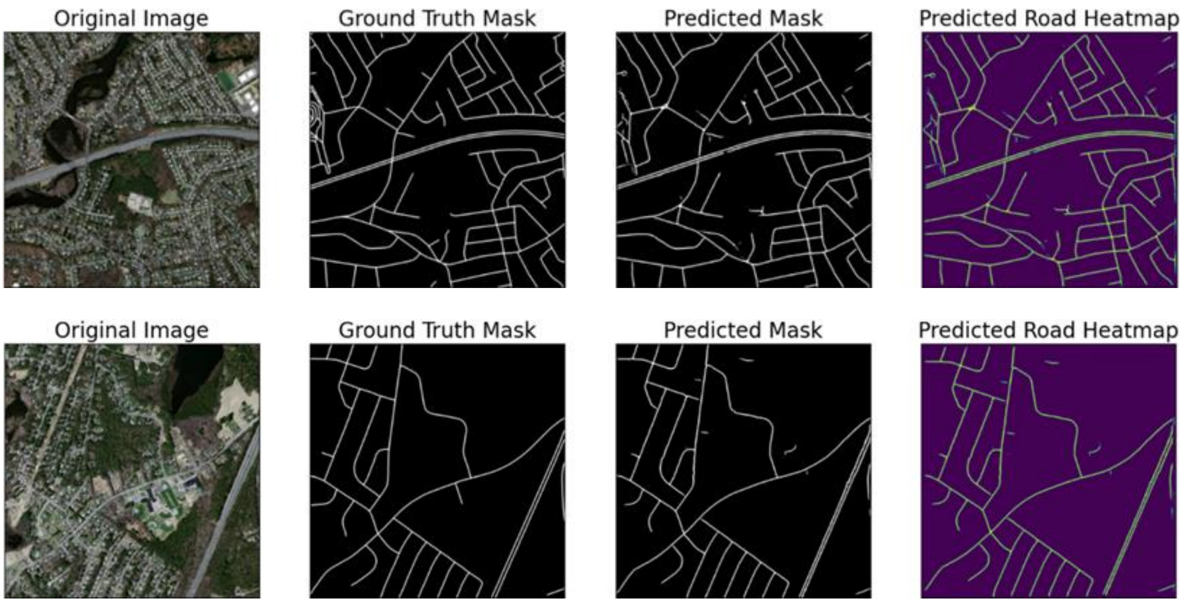

Figure 9. Model Prediction vs Ground Truths for Testing Dataset

Figure 7 shows the plot of the IoU score versus epochs for both training and validation sets. The IoU score increased during the initial epochs of training, reaching a peak of 0.942 at epoch 39 for the training set and 0.891 at epoch 39 for the validation set. After these peaks, the IoU score remained relatively constant for the rest of the training process. And Figure 8 shows the plot of the loss versus epochs for both training and validation sets. The loss decreased during the initial epochs of training, reaching a minimum of 0.031 at epoch 48 for the training set and 0.061 at epoch 48 for the validation set. After these minima, the loss remained relatively constant for the rest of the training process. In addition, the model was tested on an additional set of 50 images that were not used for training or validation. The IoU score achieved on this test set was 0.91.

Moreover, Figure 9 represents the model performance on images it has not been trained on. The model did perform well in urban areas by predicting the roads that are not even in the ground truths of the images. However, the model performance in rural areas where roads were thin and dusty is average when compared to urban areas with a lot of connected roads.

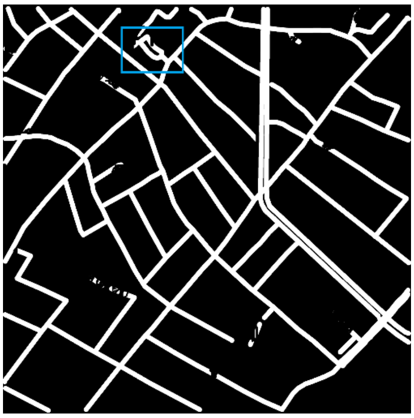

Figure 10. 2008 Predicted Mask

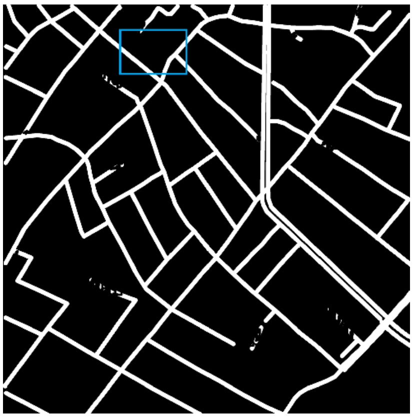

Figure 11. 2022 Predicted Mask

Figure 12. 2008 Vectorized output

Figure 13. 2022 Vectorized output

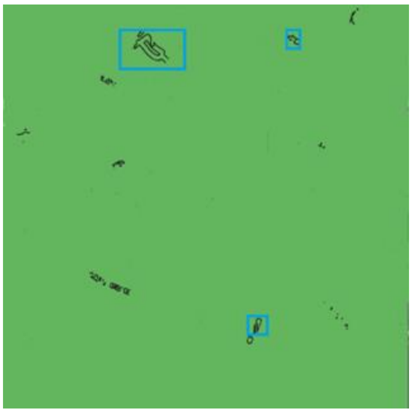

Figure 14. Change Detection Output

The outputs of the semantic segmentation for the Massachusetts images taken from different time periods were used for change detection in road networks. Figure 10 and Figure 11 represent the labels predicted by the U-Net model. These outputs were converted into vector format using the GDAL Polygonize tool. Figures 12 and 13 shows the vectorized outputs of the Massachusetts satellite images of the year 2008 and 2022 respectively. The change detection was done using the symmetrical difference tool in qgis and the output of the change detection shows changes in roads in the time periods between 2008 and 2022.

Figure 14 illustrates the identified changes in road networks of Massachusetts images taken from google earth with different time periods of 2008 and 2022. The blue rectangles are added manually to highlight the changes while remaining being uncertainties in the change detection.

4. CONCLUSION

In the proposed study, The U-Net architecture is applied to perform semantic segmentation on the Massachusetts Road Dataset and evaluated its performance on a subset of the dataset. The results showed that the U-Net architecture is a reliable model for using it for change detection, achieving high IoU scores during training, validation, and testing on a subset of the dataset. The outputs predicted by the model for the satellite images of two different time periods are used to perform change detection in road networks.

The results of the change detection showed promising results, indicating that the U-Net architecture can detect changes in road networks over time. This has important implications for urban planning and infrastructure management. However, the analysis also showed that the model's performance on rural roads was poor, suggesting that additional training data or modifications are needed to improve its performance on this subset of the dataset.

In summary, this study demonstrates the potential of the U-Net architecture for semantic segmentation and change detection in road networks. Future work could focus on improving the model's performance on rural roads, as well as evaluating its performance on other datasets.

5. FUTURE WORK

In terms of future research directions, several avenues hold promise for enhancing and extending the proposed methodology. One such avenue involves the exploration of multi-class segmentation, which would enable the categorization of various road types and urban features, thereby providing a more nuanced understanding of the evolving urban landscape. Additionally, incorporating advanced temporal analysis techniques could yield deeper insights into the dynamics of road network changes over time, moving beyond simple temporal comparisons. Integrating data from diverse sources, such as traffic flow patterns and land use data, could enhance the accuracy of change detection and contribute to a more comprehensive understanding of the underlying driving factors. Contextual information from the surroundings could be leveraged to refine semantic segmentation results, enhancing the precision of identifying road areas. An automated alert system that triggers notifications in response to significant road network changes could facilitate timely decision-making and urban planning. Extending the methodology to incorporate three-dimensional (3D) data could provide a more holistic assessment of road network changes, including variations in elevation and volume. Moreover, combining machine learning techniques with spatial-temporal models could enable predictive analysis, assisting in proactive urban planning efforts. Exploring the feasibility of real-time change detection using streaming satellite or aerial imagery holds potential for continuous monitoring and rapid response to emerging urban developments. The method's adaptability to other urban monitoring domains, like vegetation change or land cover alterations, presents an opportunity for broader applications. Finally, the development of a user-friendly software interface could make the methodology accessible and practical for urban planners, policymakers, and researchers without specialized geospatial expertise.

REFERENCES

[1] Z. Zhang, Q. Liu and Y. Wang, "Road Extraction by Deep Residual U-Net," in IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 5, pp. 749-753, May 2018, doi: 10.1109/LGRS.2018.2802944.

[2] M. Torma, P. Harma and E. Jarvenpaa, "Change detection using spatial data problems and challenges," 2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 2007, pp. 1947-1950, doi: 10.1109/IGARSS.2007.4423208.

[3] L. Zhu, A. Erving, K. Koistinen, M. Nuikka, H. Junnilainen, N. Heiska, and H. Haggrén, "Georeferencing multi-temporal and multi-scale imagery in photogrammetry," in ISPRS 2008 Beijing, The XXI Congress, The International Society for Photogrammetry and Remote Sensing, 3.-11.7.2008, Beijing China.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, N. Navab, J. Hornegger, W. Wells, and A. Frangi, Eds., vol. 9351, Springer, Cham, 2015, doi: 10.1007/978-3-319-24574-4_28.

[5] Y. Wu, "Raster, Vector, and Automated Raster-to-Vector Conversion," in Moving Theory into Practice: Digital Imaging for Libraries and Archives, A. R. Kenney and O. Y. Rieger, Eds., 2000.

[6] D. R. Panuju, D. J. Paull, and A. L. Griffin, “Change Detection Techniques Based on Multispectral Images for Investigating Land Cover Dynamics,” Remote Sensing, vol. 12, no. 11, p. 1781, Jun. 2020, doi: 10.3390/rs12111781.

[7] M. Sahu and A. Ohri, "Vector map generation from aerial imagery using deep learning," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 4, pp. 157-162, 2019.

[8] J. Botelho, S. C. P. Costa, J. G. Ribeiro, and C. M. Souza, “Mapping Roads in the Brazilian Amazon with Artificial Intelligence and Sentinel-2,” Remote Sensing, vol. 14, no. 15, p. 3625, Jul. 2022, doi: 10.3390/rs14153625.

[9] V. Hnatushenko, P. Kogut, and M. Uvarov, "Variational approach for rigid co-registration of optical/SAR satellite images in agricultural areas," Journal of Computational and Applied Mathematics, vol. 400, p. 113742, 2022, ISSN 0377-0427, doi: 10.1016/j.cam.2021.113742.

[10] M. Tahoun, A. El Rahman Shabayek, R. Reulke, and A.E. Hassanien, "Co-registration of satellite images based on invariant local features," in Intelligent Systems'2014, D. Filev et al., Eds., vol. 323, Springer, Cham, 2015, doi: 10.1007/978-3-319-11310-4_56.

[11] A. Stumpf, D. Michéa, and J.-P. Malet, “Improved Co-Registration of Sentinel-2 and Landsat-8 Imagery for Earth Surface Motion Measurements,” Remote Sensing, vol. 10, no. 2, p. 160, Jan. 2018, doi: 10.3390/rs10020160.

[12] Z. Chen, C. Wang, J. Li, N. Xie, Y. Han and J. Du, "Reconstruction Bias U-Net for Road Extraction From Optical Remote Sensing Images," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 2284-2294, 2021, doi: 10.1109/JSTARS.2021.3053603.

[13] R. Lian and L. Huang, "DeepWindow: Sliding Window Based on Deep Learning for Road Extraction From Remote Sensing Images," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 1905-1916, 2020, doi: 10.1109/JSTARS.2020.2983788.

[14] A. Abdollahi, B. Pradhan, N. Shukla, S. Chakraborty, and A. Alamri, “Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction: A State-Of-The-Art Review,” Remote Sensing, vol. 12, no. 9, p. 1444, May 2020, doi: 10.3390/rs12091444.

[15] H. Ostadabbas, H. Weippert, and F.-J. Behr, "Using the synergy of QField for collecting data on-site and QGIS for interactive map creation by ALKIS® data extraction and implementation in PostgreSQL for urban planning processes," IEEE International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B4-2020, pp. 679-683, 2020, doi: 10.5194/isprs-archives-XLIII-B4-2020-679-2020.

[16] M. Terrone, P. Piana, G. Paliaga, M. D’Orazi, and F. Faccini, “Coupling Historical Maps and LiDAR Data to Identify Man-Made Landforms in Urban Areas,” ISPRS International Journal of Geo-Information, vol. 10, no. 5, p. 349, May 2021, doi: 10.3390/ijgi10050349.

[17] F. Dahle, K. Arroyo Ohori, G. Agugiaro, and S. Briels, "Automatic change detection of digital maps using aerial images and point clouds," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B2-2021, pp. 457-464, 2021.

[18] A. Tassi and M. Vizzari, “Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms,” Remote Sensing, vol. 12, no. 22, p. 3776, Nov. 2020, doi: 10.3390/rs12223776.

[19] D. Scheffler, A. Hollstein, H. Diedrich, K. Segl, and P. Hostert, “AROSICS: An Automated and Robust Open-Source Image Co-Registration Software for Multi-Sensor Satellite Data,” Remote Sensing, vol. 9, no. 7, p. 676, Jul. 2017, doi: 10.3390/rs9070676.

[20] S. Saha, F. Bovolo and L. Bruzzone, "Unsupervised Deep Change Vector Analysis for Multiple-Change Detection in VHR Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 6, pp. 3677-3693, June 2019, doi: 10.1109/TGRS.2018.2886643.

[21] K. Lim, D. Jin and C. -S. Kim, "Change Detection in High Resolution Satellite Images Using an Ensemble of Convolutional Neural Networks," 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Honolulu, HI, USA, 2018, pp. 509-515, doi: 10.23919/APSIPA.2018.8659603.

[22] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, N. Navab, J. Hornegger, W. Wells, and A. Frangi, Eds., vol. 9351, Springer, Cham, 2015, pp. 234-241, doi: 10.1007/978-3-319-24574-4_28.

[23] M. Sobhana, G. S. D. Kumar, P. Pakkiru and Y. Tejaswi, "Extraction of Roadways from Aerial Images using U-Net Backbones," 2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/CSITSS57437.2022.10026364.

[24] S. Wang, H. Yang, Q. Wu, Z. Zheng, Y. Wu, and J. Li, “An Improved Method for Road Extraction from High-Resolution Remote-Sensing Images that Enhances Boundary Information,” Sensors, vol. 20, no. 7, p. 2064, Apr. 2020, doi: 10.3390/s20072064.

[25] H. Chen, Z. Qi and Z. Shi, "Remote Sensing Image Change Detection With Transformers," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-14, 2022, Art no. 5607514, doi: 10.1109/TGRS.2021.3095166.

[26] M. D. Hossain and D. Chen, "Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 150, pp. 115-134, 2019, doi: 10.1016/j.isprsjprs.2019.02.009.

[27] H. Chen, et al., "Change detection in multi-temporal VHR images based on deep Siamese multi-scale convolutional networks," arXiv preprint arXiv:1906.11479, 2019, doi: 10.48550/arXiv.1906.11479.

[28] X. Liu, Z. Deng, and Y. Yang, "Recent progress in semantic image segmentation," Artificial Intelligence Review, vol. 52, pp. 1089-1106, 2019, doi: 10.1007/s10462-018-9641-3.

[29] E. Brewer, J. Lin, P. Kemper, J. Hennin, and D. Runfola, "Predicting road quality using high resolution satellite imagery: A transfer learning approach," PLoS ONE, vol. 16, no. 7, pp. e0253370, 2021, doi: 10.1371/journal.pone.0253370.

[30] Y. Ren, Y. Yu, and H. Guan, “DA-CapsUNet: A Dual-Attention Capsule U-Net for Road Extraction from Remote Sensing Imagery,” Remote Sensing, vol. 12, no. 18, p. 2866, Sep. 2020, doi: 10.3390/rs12182866.

[31] V. Yerram et al., “Extraction and Calculation of Roadway Area from Satellite Images Using Improved Deep Learning Model and Post-Processing,” Journal of Imaging, vol. 8, no. 5, p. 124, Apr. 2022, doi: 10.3390/jimaging8050124.

[32] A. Shabbir, N. Ali, J. Ahmed, B. Zafar, A. Rasheed, M. Sajid, A. Ahmed, and S. H. Dar, "Satellite and scene image classification based on transfer learning and fine-tuning of ResNet50," Mathematical Problems in Engineering, vol. 2021, pp. 5843816, 2021, doi: 10.1155/2021/5843816.

[33] Y. Y. Chiang and C. A. Knoblock, "A general approach for extracting road vector data from raster maps," International Journal on Document Analysis and Recognition (IJDAR), vol. 16, pp. 55-81, 2013, doi: 10.1007/s10032-011-0177-1.

[34] L. -A. Tran and M. -H. Le, "Robust U-Net-based Road Lane Markings Detection for Autonomous Driving," 2019 International Conference on System Science and Engineering (ICSSE), Dong Hoi, Vietnam, 2019, pp. 62-66, doi: 10.1109/ICSSE.2019.8823532.

[35] M. Wang, Z. Chen, L. Mu, and X. Zhang, "Road network structure and ride-sharing accessibility: A network science perspective," Computers, Environment and Urban Systems, vol. 80, pp. 101430, 2020, doi: 10.1016/j.compenvurbsys.2019.101430.

[36] J. A. Gómez et al., “Analyzing the Spatiotemporal Uncertainty in Urbanization Predictions,” Remote Sensing, vol. 13, no. 3, p. 512, Feb. 2021, doi: 10.3390/rs13030512.

BIOGRAPHY OF AUTHORS

Dr. M Sobhana is currently working as Sr. Assistant Professor, Department of Computer Science and Engineering, V R Siddhartha Engineering College, Vijay-awada. She received Ph.D. degree in Computer Science and Engineering in 2018 from Krishna University. She has 16 years of teaching experience. Her research in-terests lie in areas such as Artificial Intelligence, Machine Learning, Data Analytics, Cyber Security and Software Engineering. She published 35 papers in National and Inter- national journals and also published 7 patents.

Gudapati Satya Dinesh Kumar is currently a under graduation student at V R Siddhartha Engineering College, Vijayawada. His research aligns in the fields of the Internet of Things, Machine Learning and Artificial Intelligence.

Yarramreddy Tejaswi is currently a under graduation student at V R Siddhartha Engineering College, Vijayawada. Her research aligns in the fields of Data Science, Machine Learning, Internet of Things and Cloud Computing.

Pavithra Pakkiru is currently a under graduation student at V R Siddhartha Engineering College, Vijayawada. Her research aligns in the fields of Data Science, Machine Learning and Computational Creativity.