



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 6, June 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Automatic Road Extraction from Remotely Sensed Images in Urban Areas

Gauravi Wakchaure¹, Prajakta Vetel², Ruchika Sinare³, Pallavi Wakchaure⁴

U.G. Student, Department of Computer Engineering, Amrutvahini College of Engineering, Sangamner, Ahmednagar, India¹

U.G. Student, Department of Computer Engineering, Amrutvahini College of Engineering, Sangamner, Ahmednagar, India²

U.G. Student, Department of Computer Engineering, Amrutvahini College of Engineering, Sangamner, Ahmednagar, India³

U.G. Student, Department of Computer Engineering, Amrutvahini College of Engineering, Sangamner, Ahmednagar, India⁴

ABSTRACT: Automatically extracting road networks from Very High Resolution (VHR) images in urban areas is a classical task. This paper represents a novel method for extracting road networks from VHR images in complex urban areas. We develop U-net architecture to extract the road networks and predict the road surface and centerlines from VHR images. The developed model can train fewer dataset but gives the best result for that dataset. In road net extraction, it solves the following issues: 1) occlusions occurred on the road surface; 2) shadows of trees and buildings 3) training the dataset with the help of convolution layers in U-net architecture 4) prediction of road surface and centerlines improves model in multitasking. 5) Supervised learning with the U-net architecture gives the more accurate results for road net extraction as compared to previous extraction methods. Dataset consisting of 500VHR images with the ground truths are provided to give the best performance for the proposed model. Optimizers used are SGD, adam and Adamax to check the accuracy for proposed model.

KEYWORDS:U-net, Convolution layer, Max pooling, Image segmentation, Centerline extraction, Road network extraction.

I. INTRODUCTION

As in an urban area, we found the road network is complex and difficult to extract as the issues occurring related to occlusions, shadows of trees and buildings. So to solve these issues and extract the road network in the area we have introduced a classical method. In the proposed model, U-net architecture is applied for training the dataset images more efficiently. This architecture can work for very few dataset giving the most accurate results for dataset provided. Generally road network extraction involves surface segmentation and centerline detection. While extraction, it faces some vision problems of occlusion and shadows on the road. So we have developed the model using U-net to overcome these vision problems and extract the road network in the complex areas.

In the model, U-net consist of two paths: 1) Contracting path; 2) Expanding path. Contracting path captures the object to be predicted and expanding path gives the localization of the prediction in the given image. This architecture consist of 23 convolution layers which train the cropped and grayscale dataset images and the surface segmentation and centerline detection is done with the highest accuracy as compared to the methodologies followed before. We have done the literature survey for the previously used methods, which are proved better. In that methods, the accuracy issues and complexity of the model were the most faced drawbacks. This model gives the best accuracy under Adamax optimizer for detection and reduces the complexity.



II. RELATED WORK

The purpose of this survey is to study the methodology used previously and find out the drawbacks of the system. So that the new model should be able to solve the issues and drawbacks found in that system and give the improved model for the detection.

W. Shi, Z. Miao, and J. Debayle [1] proposed a method to detect the road network in urban areas and presented an integrated method to extract urban main-road centerlines. The model uses multispectral high resolution images for training and testing. Here local linear kernel smoothing regression is used to extract the smooth road centerlines. Road network are generated using tensor voting. This model gives the results at its best but fails to handle the complicated road junctions. Another model was proposed by R. Alshehhi, Rasha, Prashanth Reddy Marpu, Wei Lee Woon, and Mauro Dalla Mura [2] to extract the man-made objects (like roads and buildings) from remotely sensed imagery in many urban applications. The model uses the method of a single patch-based Convolutional Neural Network (CNN) architecture for extraction of roads and buildings from high-resolution remote sensing data. This model extracts the road in area and gives the best results for the plain road surfaces but fails to detect the surfaces with the issues like occlusions and shadows. Also the method fails to detect the broken or irregular road boundaries.

E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez [3] proposed with some modifications another classical method using systematic review of CNN model for large-scale image dataset was introduced. The model consist of convolution layers to train the large-scale images and perform the road extraction in remote sensing images by the current deep learning techniques applied to remote sensing datasets for road extraction and semantic segmentation. The method proved better in the extraction. Later on, Marmanis, Dimitrios, Konrad Schindler, Jan Dirk Wegner [4] has given an approach towards an end-to-end trainable deep convolutional neural network (DCNN) for semantic segmentation with built-in awareness of semantically meaningful boundaries was introduced. Author has represented a technique for the prediction, the method added boundary detection to SEGNET encoder-decoder architecture. Also boundary detection in FCN-type model is done which sets high-end classifier. The model used is more complex. The size of the model makes an issue as it has to perform many operations and this makes the model more complicated.

III. METHODOLOGY

This section consists of the detailed architecture of our proposed method. We have used U-net architecture [5] introduced by the authors Olaf Ronneberger, Philipp Fischer, and Thomas Brox for biomedical image segmentation. The layers included in the architecture which improved the accuracy of detection have been explored further.

1) U-net architecture

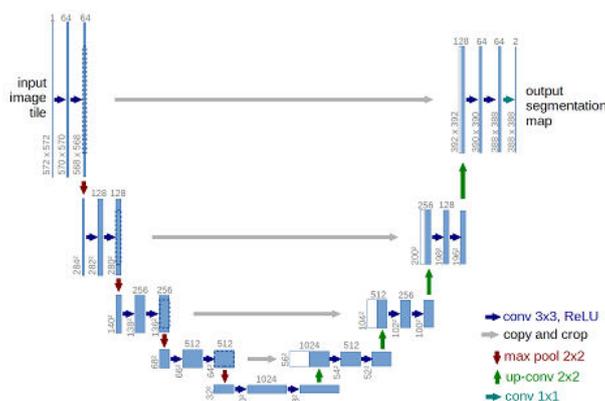


Fig. 1. U-net architecture. This architecture consists of multi-channel feature maps represented by blue boxes and copied feature maps are represented by white boxes. Also arrows with the different colors performs different operations.



In this section, we are formulating the road surface as a binary image labeling, where “1” and “0” refer to road surface and non-road surface pixels in the input image. It performs two tasks: convolution and deconvolution in contracting and expansive path. The architecture consists of 23 convolutional layers.

As shown in fig.1, the contracting path of the architecture which includes the convolutional process is symmetric to the expansive path which consist of up sampling technique. The blue boxes containing the input images are connected to each other by the different arrows performing different tasks. In encoder, the convolution process followed by rectified linear unit is performed. The red arrows in the architecture are the max pooling operation which halves the image size using padding. This is referred as down sampling of image. As the image size is reduced, the process increases the depth of the image. For the next expansive path, the up sampling of the image is done by the deconvolutional layer. Here the reduced image is gone through the up sampling technique and it regain its size. At the end performing convolutional layer, the output is extracted in original size of input image.

IV. EXPERIMENT

1) Implementation

The U-net training is performed on the kera-tensorflow platform which is publically available. The parameters used for U-net are: input image of size (256×256), the ground truth of size (256×256), batch size (10), with number of epochs (90). We have used Adam, SGD and Adamax optimizers for performance checking. We have included the results according to the optimizers used and improved accuracy. We performed U-net training on the dataset having (256×256) dimension. The experiment is carried out using Intel(R) Core(TM) i5-7200U CPU at 2.50 GHz with 8 GB RAM.

2) Data

The dataset we have used is Ottawa dataset from google earth. These images are with the spatial 0.21 m resolution. For processing, the images in dataset with the dimensions around 4000 × 3000 are cropped to the dimensions 256 × 256 using cropping method. With this, we can extract the images with maximum focus on the extraction part of image. In Fig.2, we can see the original image is cropped and extracted the image with 256 × 256 dimensions for the selected region and the surface and centerline ground truths as per the cropped image. Similarly, we have cropped the original data and created 650 images with their ground truths. These dataset is divided for the training purpose as 500 images for training, 100 images for validation and 50 images for testing.

In Fig. 3, we have shown some samples from the dataset including ground truths of road surface and centerline.

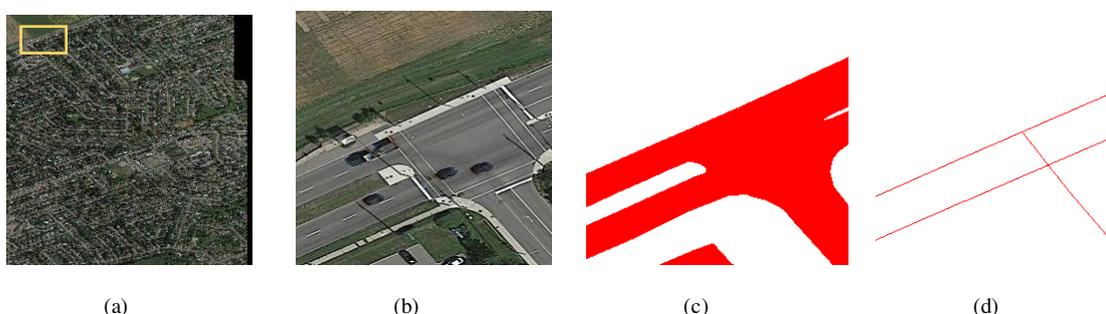


Fig. 2. Dataset representation which consist: a) a VHR image with dimensions 4007 × 5888 representing area to containing road network b) the cropped image with the dimensions 256 × 256 illustrated from the original image c) the road surface ground truth of the cropped image d) centerline ground truth.



Fig. 3. Sample images from the dataset after cropping the dataset into 256×256 dimensions for a) input images after cropping into 256×256 . b) ground truth for road surface segmentation with dimension 256×256 . c) ground truths for centerline with dimensions 256×256 .

3) Optimizers

Optimizers are used for the changes in the attributes like weights and learning rates. This change is performed to reduce the losses in respected architecture. In our model, we have performed U-net training on our dataset images. In that, to get the maximum accuracy and reduce the loss, we have trained the data for three optimizers: SGD, adam, and Adamax.

SGD is stochastic gradient descent, which is a method for optimizing an objective functions with suitable smoothness properties. In SGD optimizer, the update direction is dependent on current data, so its update is unstable. In SGD, learning factor, momentum are used as training parameters. Here we have taken 0.01 as learning factor and 0.9 is momentum. Nesterov momentum is by default taken as false. The model is trained for these parameters and gives the result as shown in fig. 6. As the resultant output for SGD optimizer is not satisfactory, we trained the model for adam optimizer. Adam [9] is an optimization algorithm which can be used as an alternative for stochastic gradient descent. Adam is based on two moments first-order and second-order. It can be used to update network weights in training data. Adam algorithm calculates an average of the gradient and the squared gradient. The parameters used here are beta 1, beta 2, and amsgrad. Here, beta1 contains a float value i.e. 0.9 this is decay rate for 1st moment and beta 2 also contains float value of 0.999 which is the decay rate for the 2nd moment. Amsgrad contains boolean value which is



taken as by default i.e false. The training results for the adam optimizer are shown in fig. Results obtained are much better with maximum accuracy. Adamax is the optimizer that implements the Adamax algorithm. It is a variant of Adam. Adamax is sometimes superior to adam. So we trained the dataset for the adamax optimizer. Parameters we used here are learning_rate, beta_1, beta_2, epsilon, name, **kwargs. So, we have taken learning_rate as 0.001 and 1st momentum as 0.9 and as 0.999. Epsilon is a small constant value which is 1e-07. kwargs is an keyword arguments. The results for the adamax optimizer are included in fig.

V. EXPERIMENTAL RESULTS

After training the model for three different optimizers, we perform testing on random images from the dataset. In fig. 4, we have taken random input images to test the results with ground truths for the road surface and centerline.



Fig. 4. Random input VHR images to be extracted from the dataset provided to the model. Images are complex and consist of challenges like shadows and occlusions by the trees and buildings. The content in the fig is (a) Input images for the testing purpose (b) ground truth of the road surface for respective input image (c) ground truths of the centerlines for respective road network images.

After performing the training on the images given in dataset, here are the results according to the optimizers used. In fig. 5(a) are the original input images on which testing has done. Fig. 5(b) contains the resultant output for the SGD optimizer. The output is not clear and the accuracy is poor. Here loss is maximum so the output image is not giving the maximum accuracy for this optimizer. In fig. 5(c), the output is obtained by using the adam optimizer. Here the resultant output images are much clear and the accuracy is better than the SGD optimizer. The loss of the model is reduced and performance is increased. In fig. 5(d), we have the results of the training using Adamax optimizer. Results of the adam optimizer and Adamax optimizer are not much different. Adamax can be proved better than adam for particular conditions over parameter. By this, the proposed model gives the best result using adamax optimizer.

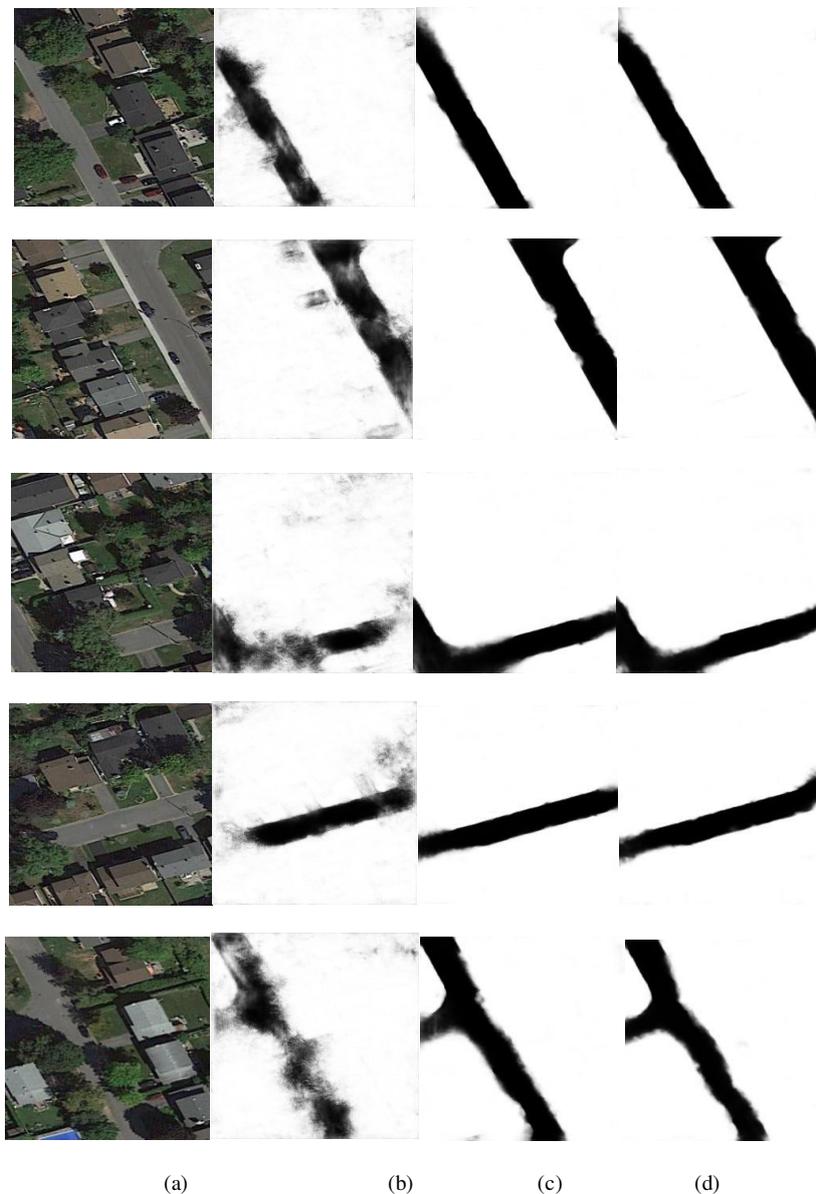


Fig. 5. Experimental results for the proposed model containing: (a) Original input image for testing. (b) Extracted road surface images using SGD optimizer. (c) Road surface extraction using Adam optimizer. (d) Extracted road surface by the Adamax optimizer.

For centerline extraction, the ground truth of the centerlines were too invisible for the training as in fig. 4(c). So, to make them visible we have performed erosion operation on them. By the changing iterations we made them broader for training purpose (fig. 6(a)). After the erosion operation, by performing the training, we got the output for these erode images in fig. 6(b). For final extraction, the output images were gone through the dilation operation. This operation is performed to make the centerlines display in original size again. Then we got the finally extracted centerlines in fig. 6(c).

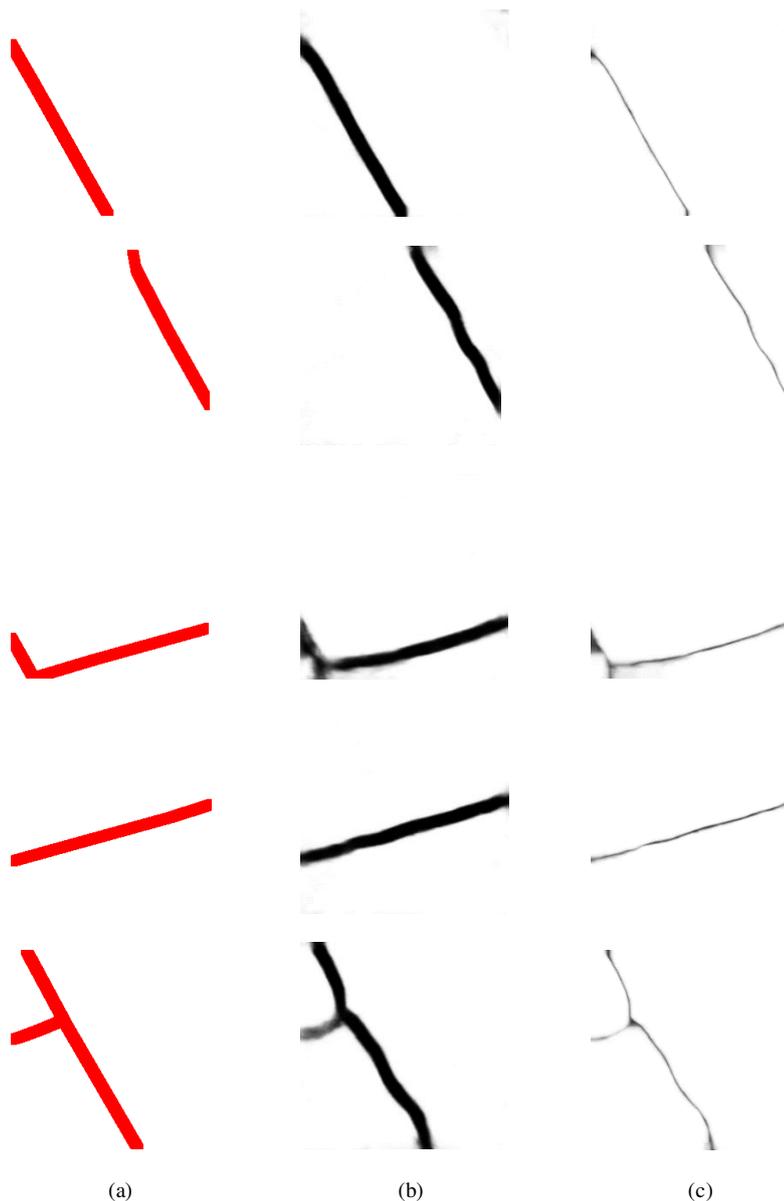


Fig. 6. The centerline ground truth images which are operated for the testing which contains: (a) centerlines after performing the erosion operation (b) output images using Adamax as an optimizer (c) dilated centerline images with original size.

VI.CONCLUSION

We have implemented an automatic road network extraction from Very High Resolution remotely sensed images using a classical method U-net [5] architecture. We have performed the tasks road surface segmentation and centerline extraction even in the extreme condition like occlusions on the road and the shadows by the trees and



buildings. The model has performed the classification on the complex image dataset provided in input. So, we have got the output in the form of extracted road surface and centerlines. We have developed this model to solve the issues related to vision and to provide the methodology which gives the maximum performance and accuracy for the extraction and avoided the complexity in the model. The U-net architecture gives best accuracy for the fewer dataset but for dataset with larger data, it is time consuming process. Also memory issue can be the problem for the GPU.

REFERENCES

- [1] W. Shi, Z. Miao, and J. Debayle, "An integrated method for urban main-road centerline extraction from optical remotely sensed imagery," *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, VOL. 52, NO. 6, JUNE 2016.
- [2] R. Alshehhi, Rasha, Prashanth Reddy Marpu, Wei Lee Woon, and Mauro Dalla Mura, "Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks" *ISPRS Journal of Photogrammetry and Remote Sensing* 130 (2017): 139-149.
- [3] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Fully convolutional neural networks for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing* (Volume:55, Issue:2, 2017).
- [4] Marmanis, Dimitrios, Konrad Schindler, Jan Dirk Wegner, Silvano Galliani, Mihai Datcu, and Uwe Stilla. "Classification with an edge: Improving semantic image segmentation with boundary detection" *ISPRS Journal of Photogrammetry and Remote Sensing* 135 (2018): 158-172.
- [5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation" *arXiv:1505.04597v1 [cs.CV]* 18 May 2015.
- [6] Y. Liu, X. H. Ming-Ming Cheng, K. Wang, and X. Bai, "Richer convolutional features for edge detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5872–5881.
- [7] Yahui Liu, Jian Yao, Member, IEEE, Xiaohu Lu, Menghan Xia, Xingbo Wang, and Yuan Liu, "RoadNet: Learning to Comprehensively Analyze Road Networks in Complex Urban Scenes from High-Resolution Remotely Sensed Images" *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, VOL. 57, NO. 4, APRIL 2019
- [8] T. Panboonyuen, K. Jitkajornwanich, S. Lawawirojwong, P. Srestasathiern, and P. Vateekul, "Road segmentation of remotely sensed images using deep convolutional neural networks with landscape metrics and conditional random fields," *Remote Sens.*, vol. 9, no. 7, pp. 2072–4292, 2017.
- [9] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2014, *arXiv preprint arXiv:1412.6980*.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

**Impact Factor:
7.512**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details