International Journal of Engineering & Technology, 7 (4.10) (2018) 1079-1084

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Road Extraction using Deep Learning

J. D. Dorathi Jayaseeli [1], D. Malathi[2]*, Gopika S[3]

[1]*Assitant Professor(S.G), CSE Department, SRM Institute of Science and Technology, Kattankulathur*
[2]*Professor, CSE Department, SRM Institute of Science and Technology, Kattankulathur*
[3]*Research Scholar, CSE Department, SRM Institute of Science and Technology, Kattankulathur*
**Corresponding author E-mail: malathi.d@ktr.srmuniv.ac.in*

Abstract

The Road extraction from aerial image, stands as a quintessential node for the development of rudimentary layers in innumerable fields. From GIS, to Unmanned Aerial vehicles, road maps pave the foundation for data accumulation. This significant process is a result of number of mechanisms devised over the years through iterative experiments and research. However, the glut of methods available often pose as a hurdle in the selection process. In this project we implement a novel approach to solve the extraction problem, by incorporating generative algorithm using conditional adversarial networks. We investigate conditional adversarial networks as a general-purpose solution to image-to-image translation problems. These networks not only learn the mapping from input image to output image, but also learn a loss function to train this mapping. This makes it possible to apply the same generic approach to problems that traditionally would require very different loss formulations. The U-Network incorporated essentially convolves and de-convolves over the generative model, thus producing a pixel to pixel image translation, the result of which is the vector road map of its corresponding aerial image. The entire model is trained on a 990 MS GPU for computational ease.

Keywords: *Remote Sensing Data, Road Extraction, Roadmaps, Features, Classification Methods, Artificial Neural Network*

1. Introduction

Image processing and Computer vision have evolved over the years to account for nuances in the pragmatic world, and extend a visual cortex to technology. This development has been a boon to aerial images. With advances such as segmentation, pruning, Principal Component Analysis (PCA), and self- learning artificial intelligence implementations such as Artificial Neural Networks (ANN), Support Vector Machine (SVM), and classification, quality of image data sets have surged immensely. However, the paramount notion that these advancements are based on unique circumstances and that their efficiency relates to the appropriate usage [11], has now vaporized into the obsolete. In order to discern the perfect extraction mechanism, it is essential that the details of road features, their singularity, and the context of their application is understood as a base.

Road Features: Image characteristics of road features are contingent on sensor type, weather and light fluctuations, spatial and spectral resolution, and ground characteristics. Such dependencies pose a hurdle in road extraction from RS data. Intuitively, a modicum of difference in these elements propagates into large variations making is difficult tofix upon environmental variables. Hence, it is essential to analyze and fixate upon the number and type of road features, for accurate road extraction.

In general, image enhancements play a vital role in the extraction process. A road in an RS image appears as elongated geometric features with slowly changed grey values. For ease of understanding, the road features in an image are summarized from four different aspects [20]. Based on their description, they can be concluded as follows:

Geometric Features: A stripe featured road is the one which possesses a near consistent width accompanied by elongated lengths. Essentially the ratio between length and width is very large.

Photometric Features: Photometric features emphasize on the distinction between road and non-road edges. The two obvious road edge lines provide a large edge gradient [7]. Meanwhile, the grey values or colours of roads remain relatively consistent, but vary from those of the neighbouring non-road areas such as trees and buildings, etc.

Topological Features: Generally, a road has intersections. The road network is not suddenly interrupted.

Functional Features: A road has specific functions in the real world. The cycle lane, truck lane, poses as functional units. In order to realize those functions, it must have some constraints or conditions.

Texture Features: Every feature in an image represents the phenomenon of uniformity which is the texture or visual feature of the image. Interestingly, they are independent of colour and intensity information. The crux of these features is to highlight the geographical distribution of the pixel grey in the adjoining region [21].

Different road features in an image have different properties for road extraction. Geometric features are directly linked to the road shapes. Photometric features manufacture and produce the grey levels or colours. Topological features and functional features are comparatively intuitive but complex for real-life application. In practice, several road extraction methods use a combination of road features over a single entity. However, due to the influence of illumination, shadow, and occlusion, the above-mentioned features contribute in erratic amounts, making it difficult to extract road from an RS image.

This project proposes road extraction from remotely sensed data using a novel generative approach with a Generative Adversarial

Network (GAN) model. The form of learning differs from discriminative/ predictive algorithm by aiming for image- image translation. Usually the learning algorithms try to find a label for the given set of features whereas GANs try to find out the features from the given label.

2. Generative Adversarial Networks

Generative Adversarial Networks (GANs) consists of two deep neural networks called the generator and the discriminators which are working exactly in opposite way. The idea of GANs were first given by Ian Good fellow.

Generative algorithm tries to generate features for a given class whereas the discriminative algorithm tries to predict the class for the given feature set.

For example, given all the features of an image, a discriminative algorithm will find out whether the image is a car or not a car. This problem can be expressed mathematically asp(y|x) where y is the class label and x is the feature set. This means that "the probability of y given x", which in turn means "the probability that an image is a car given the features it contains."

On the other hand generative algorithm will find the feature set for the given class that is in order to say an image to be a car what are all the features it should have or to say an image to be not a car what are all features it should have.

One another way to distinguish discriminative algorithm from generative algorithm is: Discriminative algorithm learns the boundary of classes whereas generative algorithm frames the distribution of individual classes.

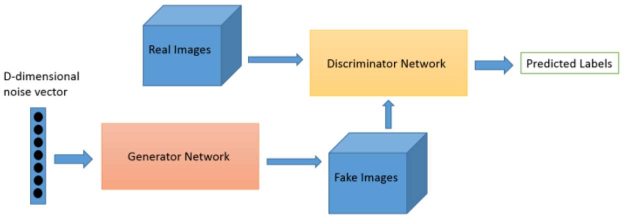


Fig. 2.1: A typical generative adversarial model

Working: The two deep neural networks called the generator and discriminator work in the opposite way. The discriminator network is actually a Convolutional Neural Network which takes as input an image and then down samples it to their features and finally predicts whether the input is real image or a fake image.

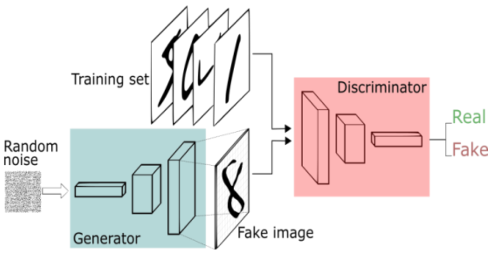


Fig.2.2: A typical GAN working

The generator is an Inverse Convolutional Neural Network so it takes as input a vector of noise and then up samples and produce a fake image. Thus both the networks become stronger by learning continuously. Thus the job of generator is to generate new fake images, whereas the job of discriminator is to decide whether the images belong to training dataset or not.

3. Conditional Adversarial Network

When the noise input to the generator is replaced with a completely different image, requesting for a generation of its mapped output, the model is named as conditional adversarialnetwork.[30] The entirety of working remains the same, however, Pz(z) now represents a whole new image, and G(z) is now its vector road map.

Formulation: Conditional Adversarial Network (CGAN) typically utilizes the probability distributions where the following are regarded as the standards.

Pdata(x) : Distribution of data (vector road maps)
X : A sample from the distribution of data
Pz(z) :A new image, given to a generator(Aerial images)
Z : A sample from the distribution new images
G(z) :Generator Network
D(x) :Discriminator Network

The training occurs as a one on one match against the two networks. The generator tries to minimize, whereas the discriminator tries to maximize the function V.The first term forms the entropy of data for the discriminator, Pdata(x).It aims to maximise the value of output to 1.

$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

The second term forms the entropy of aerial images Pz(z) that the generator generates. Here, the discriminator aims to maximise the value of output to 0. In other words, the generator tries to generate output such that the log probability of its generated outputs deemed as fake, is 0.

Thus, they train over an iterative approach, pitting against each other, where G tries to minimize, and D tries to maximize.

3.1. Advantages of conditional AN

1. CGANs learn a structured loss. Structured losses penalize the joint configuration of the output. Our conditional GAN is different in that the loss is learned, and can, in theory, penalize any possible structure that differs between output and target.
2. Our framework differs in that nothing is application-specific. This makes our setup considerably simpler than most others.
3. Significantly faster, and computationally convenient
4. Larger, meaningful goal for easier implementation
5. Reduced computational complexity
6. Can utilize full efficiency of batch- normalization
7. Can extract features which are microscopic in the image, i.e invisible to the naked eye in an aerial image.
8. No blur images are generated

4. Conditional Adversarial Network Architecture

4.1. Generator Architecture

The generator architecture typically follows a U- Network. U-Network consists of convolutions followed by de-convolutions.

The number of layers involved in the U-net is 28 including the input and output as seen in Fig 4.1.

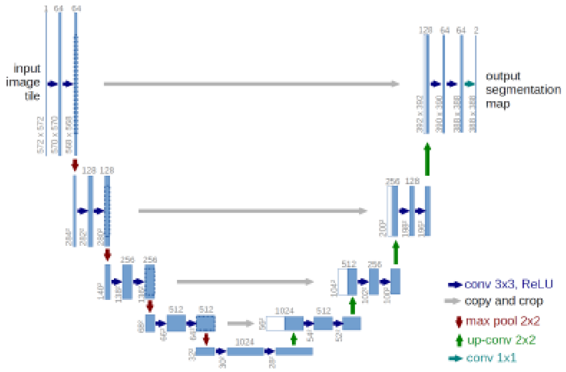


Fig. 4.1: U-Network architecture used for the generator

4.2. Discriminator Architecture

Termed as a Patch-GAN module, the discriminator is created a patch identifier. [31]The discriminator is made up of an inbuilt convolutional network, which classifies if each patch of N*N pixels in an image belongs to the fake class or real. Averaging the output finally produces the output of the discriminator. The smaller size of the filer helps increase computational time, while producing better quality images

4.3. Entity Relationship Diagram

Fig 4.2 depicts the ER diagram of the project module

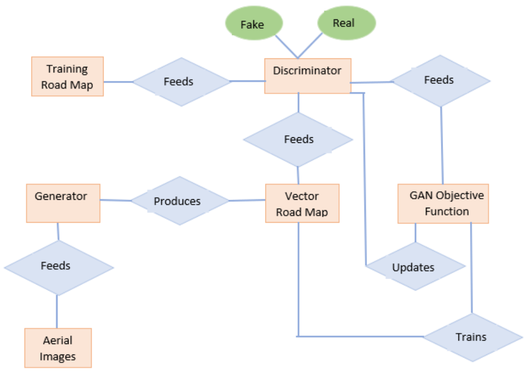


Fig. 4.2: E-R diagram for the model

5. Dataset Specifications

1. The dataset was obtained from University of Toronto's website.
2. Named: Road and Building dataset
3. The dataset contains aerial images of Massachusetts road
4. The entire dataset contains aerial images and their corresponding target vector road maps.
5. Aerial images were originally of 1500* 1500 *3 pixels of the format TIFF as shown in Fig 5.1.
6. Corresponding road maps of 1500*1500*1 of the format TIF is shown in Fig 5.2.
7. Train images: 1108 input and target, Test images: 29 input and target, Validation images: 14 input and target
8. Dataset modified for experiment: 105 * 105 * 3 for aerial images in JPEG format, 105 *105*1 for road maps in JPEG format.

Fig.5.1: Sample Aerial Image

Fig. 5.2: Corresponding Vector Road Map

6. Training

Fig 6.1, 6.2, 6.3, 6.4, 6.5 represent screenshots of training instances spread across 25000 iterations.

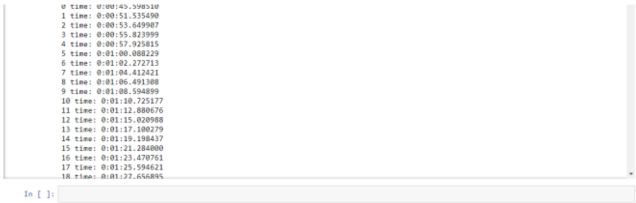


Fig. 6.1: Training snapshot of epochs 0 to 18

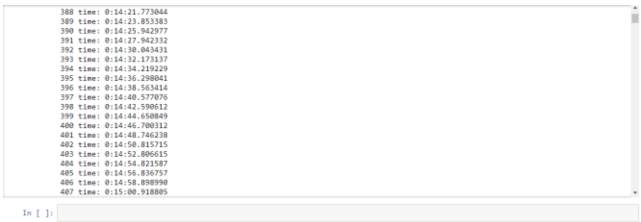


Fig.6.2: Training snapshot of epochs 388 to 407

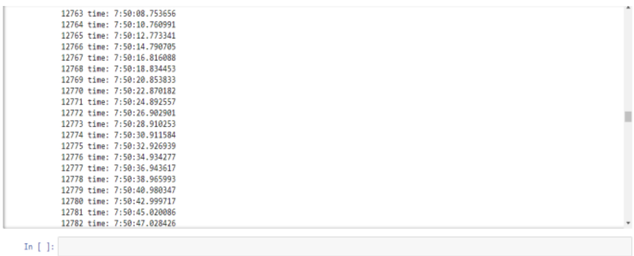
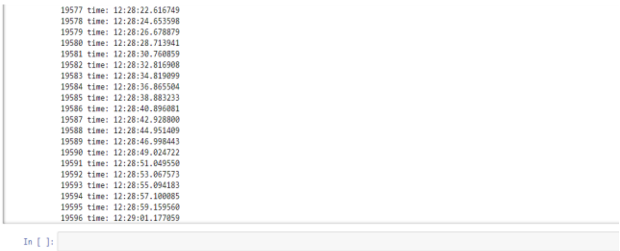


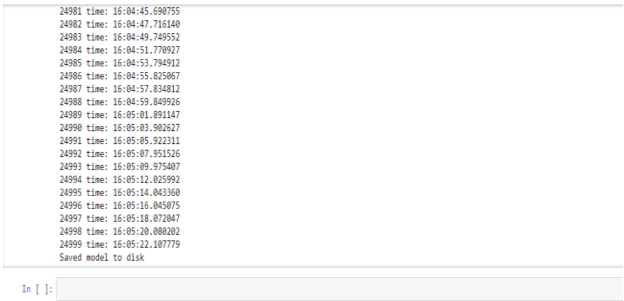
Fig.6.3: Training snapshot of epochs 12763 to 12782



```
19577 time: 12:28:22.616749
19578 time: 12:28:24.653598
19579 time: 12:28:26.678879
19580 time: 12:28:28.713941
19581 time: 12:28:30.760859
19582 time: 12:28:32.816908
19583 time: 12:28:34.819099
19584 time: 12:28:36.865504
19585 time: 12:28:38.883233
19586 time: 12:28:40.896081
19587 time: 12:28:42.928800
19588 time: 12:28:44.951409
19589 time: 12:28:46.998443
19590 time: 12:28:49.024722
19591 time: 12:28:51.049550
19592 time: 12:28:53.067573
19593 time: 12:28:55.094183
19594 time: 12:28:57.100085
19595 time: 12:28:59.159560
19596 time: 12:29:01.177059

In [ ]:
```

Fig.6.4: Training snapshot of epochs 19577 to 19596



```
24981 time: 16:04:45.690755
24982 time: 16:04:47.716140
24983 time: 16:04:49.749552
24984 time: 16:04:51.770927
24985 time: 16:04:53.794912
24986 time: 16:04:55.825067
24987 time: 16:04:57.834812
24988 time: 16:04:59.849926
24989 time: 16:05:01.891147
24990 time: 16:05:03.902627
24991 time: 16:05:05.922311
24992 time: 16:05:07.951526
24993 time: 16:05:09.975407
24994 time: 16:05:12.025992
24995 time: 16:05:14.043360
24996 time: 16:05:16.045075
24997 time: 16:05:18.072047
24998 time: 16:05:20.080202
24999 time: 16:05:22.107779
Saved model to disk

In [ ]:
```

Fig.6.5: Training snapshot of epochs 24981 to 24999

7. Outputs

The training of generator and discriminator occur side by side. By iterating for 25000 epochs, the model improves with each epoch. Fig 7.1, 7.2, 7.3, 7.4, 7.5, 7.6 and 7.7 are the generated outputs for random images from the test file for specific epochs. Each of the images contains three random aerial images from the test data set named as condition, their corresponding generated output road map under the name generated, and their original road map under the name original.

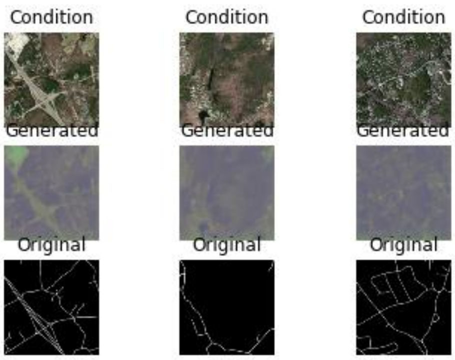


Fig. 7.1: Three aerial images, their generated output from the model, and their corresponding manual vector map outline for epoch 0

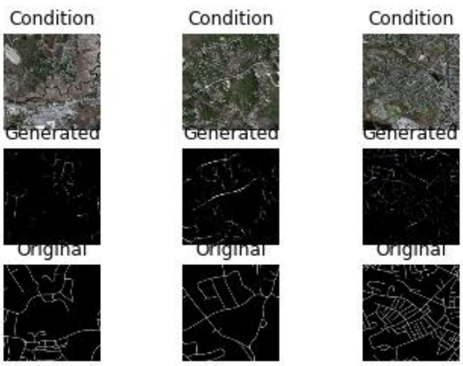


Fig. 7.2: Three aerial images, their generated output from the model, and their corresponding manual vector map outline for epoch 3000

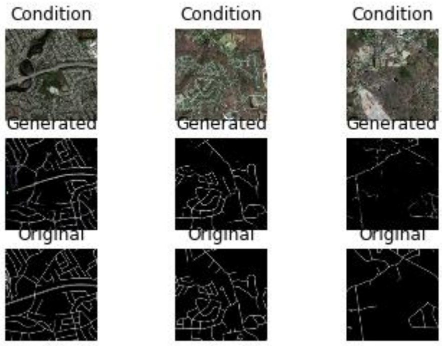


Fig. 7.3: Three aerial images, their generated output from the model, and their corresponding manual vector map outline for epoch 6000

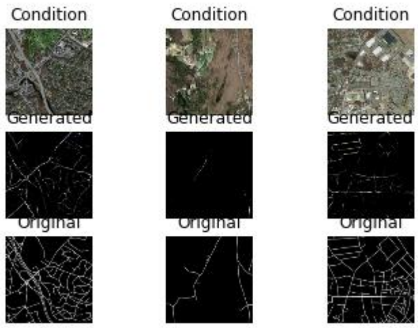


Fig. 7.4: Three aerial images, their generated output from the model, and their corresponding manual vector map outline for epoch 9000

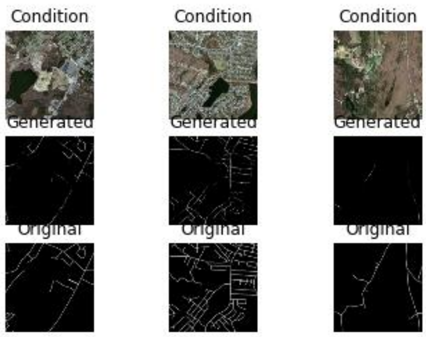


Fig. 7.5: Three aerial images, their generated output from the model, and their corresponding manual vector map outline for epoch 12000

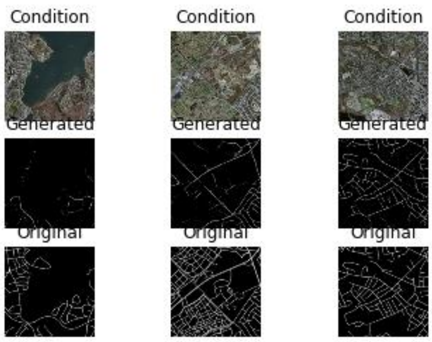


Figure 7.6: Three aerial images, their generated output from the model, and their corresponding manual vector map outline for epoch 18000

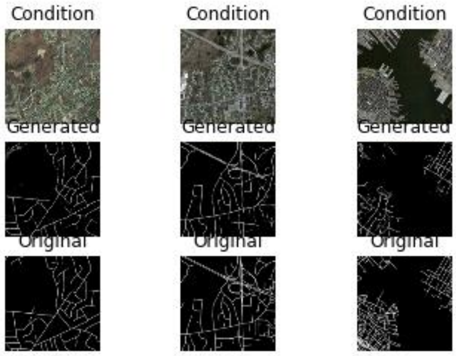


Figure 7.7: Three aerial images, their generated output from the model, and their corresponding manual vector map outline for epoch 24999

8. Performance metrics and outcomes

The model efficiently generates vector road maps for a new aerial image. Fig 8.1 and 8.2 represent the aerial image and its corresponding vector road map.

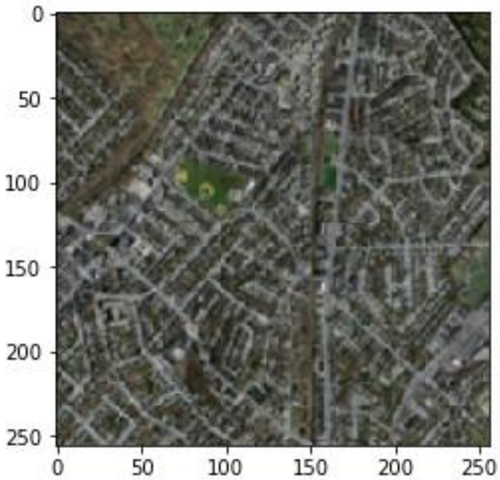


Figure 8.1: Aerial Image

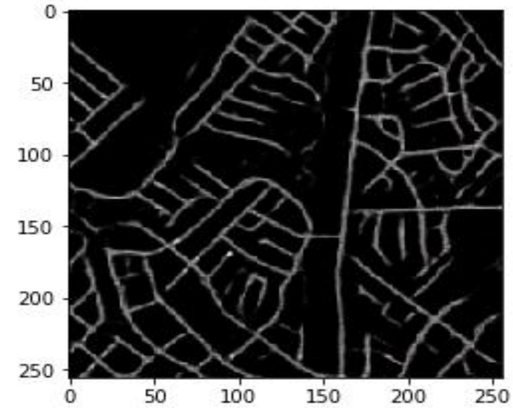


Figure 8.2: Vector Road Map

Two accuracy measures are used to estimate image-image pixel similarities for the generated vector road maps. Table 1, represents modelled accuracy in terms of SSIM metric, and inverse fakeness percentage, for 7 novel random aerial images and their corresponding road map generated**.**

SSIM is a traditional similarity metric used for comparing one image with another. On the other hand, the discriminator of the network generates a a fakeness percentage to qualify as a road map. The inverse of fakeness translates to accuracy of the generated image, with respect to its road map. This is the second metric used to measure the performance of the model.

Table 1: Accuracy measures for the model over predicted aerial road maps from real time data.

Prediction Attempt	Average SSIM Similarity Percentage for 3 Images (Complete Model)	Corresponding Accuracy Percentage (Discriminator)
1	94.6	95.3
2	93.7	95.3
3	98.1	96.2
4	96.0	96.3
5	94.8	95.4
6	99.3	96.7
7	95.6	95.6

Analysis of the performance for seven random sampling from the real time image dataset, for the saved model indicates an average of 96 % as similarity percentage among the generated and manually outlined road maps. On the other hand, percentage of trueness, translated to accuracy states that an average of 95.8 % is the accuracy of the model 's fakeness identifying capability.

Table 1 's data states that, the model has a positive implementation advantage of rendering no less than 93.7 % as similarity measure, thus proving that the bigger goal of a GAN to generate maps that are as similar as possible to the manual vector road map, proved to be successful.

9. Conclusion and Future Enhancements

The novel approach implemented surpassed the previous solutions in that, it reduced the possibility of blur images in entirety. The bigger goal proved to be a huge boon, in terms of generating clear as well as visibly similar images.

However, apart from the quality of the image as such, image production as a whole, provided generalization for varied applications. The model can be reused, and trained with different set of data, to extract varied edges and features. The model can also be turned inside out, to learn to generate an aerial image given a vector road map. Such generalizations increase its application list without much modifications required.

As a result computational complexity is also reduced. The generative algorithm solved all the problems with a panacea-tic solution. Hence, the method can be regarded as an affordable user-friendly approach for commercializing road extraction.

Future enhancements can include a post-processing step for smoothening the output vector maps. Although the model is general in purpose, and accounts for any kind of aerial image, with an advanced softening method, the vector maps can be mastered for tailoring needs.

References

[1] Abraham L., Sasikumar M., “A fuzzy-based road network extraction from degraded satellite images,” *International Conference on Advances in Computing, Communications and Informatics,* Mysore, 2013.

[2] Anil P.N., Natarajan S., “A novel approach using active contour model for semi-automatic road extraction from high-resolution satellite imagery,” *In Second International Conference on Machine Learning and Computing,* Shijiazhuang, 2010.

[3] Barzohar, M., Cooper, D.B., “The automatic finding of main roads in aerial images by using geometric-stochastic models and estimation,”. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 18, No.7, (1996), pp. 707-721.

[4] Caparrini F., E. Caporali and F. Castelli, “Neural network analysis of satellite images for land cover discrimination Mediterranean

Storms," *Proceedings of the EGS Plinius Conference held at Maratea*, Italy, 1999.
[5] Baumgartner A., Steger C., Mayer H., "Automatic road extraction based on multi-scale, grouping, and context," *Photogrammetric Engineering and Remote Sensing*, Vol.65, No. 7, (1999), pp.777-785.
[6] Ghule Swati A., and T. Rajani Mangala, "Road Network Extraction Using Support Vector Machines," *International Journal of Scientific & Engineering Research*, Vol. 6, No. 10, (October-2015), pp.521-526.
[7] Cem U., Beril S., "Road network detection using probabilistic and graph theoretical methods," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 50, No. 11, (2012), pp. 4441-4453.
[8] Chen Z., "Research on High-resolution RS Image Classification Technology," (Ph.D thesis). *Chinese Academy of Science*, Beijing. 2006.
[9] Fua P. and Leclerc Y.G., "Model driven edge detection," *Machine Vision and Application*, Vol. 3, No. 1, (1990), pp. 45- 56.
[10] Heermann P.D., Khazenie N., "Classification of multispectral remote sensing data using a backpropagation neural network," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, No.1, (1992), pp.81-88.
[11] Hinz S. and Baumgartner A, "Automatic extraction of urban road networks from multi-view aerial imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 58, No.1, (2003), pp. 83-98.
[12] Hu J., Razdan A., Femiani J.C., "Road network extraction and intersection detection from aerial images by tracking road footprints," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, No. 12, (2007), pp. 4144-4157.
[13] Isola P, Zhu JY, Zhou T, Efros A., "Image-to-image translation with conditional adversarial networks," *arXiv preprint.in CVPR* (2017).
[14] Kirthika A., Mookambiga A., "Automated road network extraction using artificial neural network," *In IEEE International Conference on Recent Trends in Information Technology*, Chennai, (2011).
[15] Hangzhou Liu, J. and Wang H.Q., "An interactive image segmentation method based on graph theory," *Journal of Electronics and Information Technology*, Vol. 8, No. 30, (2008), pp. 1973-1976.
[16] Melgani, F., Bruzzone L., "Classification of hyper-spectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 8, (2004), pp. 1778-1790.
[17] Miao Z.L., Wang B. and Shi W., "A semi-automatic method for road centerline extraction from VHR images," *IEEE Geoscience and Remote Sensing Letters*, Vol. 11, No. 11, (2014), pp.1856-1860.
[18] Mnih Volodymyr., "Machine learning for aerial image labeling" Ph.D diss., *University of Toronto* (Canada), (2013).
[19] Mokhtarzade M. and Valadanzoej M.J., "Road detection from high-resolution satellite imagery using artificial neural networks", *International Journal of Applied Earth Observation and Geoinformation*, Vol. 9, No.1, (2007), pp. 32-40.
[20] Ronneberger O, Fischer P, Brox T. U-net, "Convolutional networks for biomedical image segmentation", *In International Conference on Medical image computing and computer-assisted intervention* (2015), pp. 234-241.
[21] Simler C., "An improved road and building detector on VHR images," *In International Geoscience and Remote Sensing Symposium, Vancouver*, (2011).
[22] Tu-Ko K., "A Hybrid Road Identification System Using Image Processing Techniques and Backpropagation Neural Network," *Mississippi State University*, Starkville, (2003).
[23] Tupin F., Maitre H. and Mangin J.F., "Detection of linear features in SAR images: application to road network extraction," *IEEE Transaction on Geoscience and Remote Sensing*, Vol. 36, No. 2, (1998), pp. 434-453.
[24] Wang J.H., Qin Q.M. and Yang X., "Automated road extraction from multi-resolution images using spectral information and texture," *In International Geoscience and Remote Sensing Symposium, Quebec City*, (2014).
[25] Wang M. and Luo J.C., "Extracting roads based on Gauss Markov random field texture model and support vector machine from high-resolution RS image," *IEEE Transaction on Geoscience and Remote Sensing*, Vol. 9, No. 3, (2005), pp. 271-276.
[26] Wang Y., Zheng Q., "Recognition of roads and bridges in SAR Images," *Pattern Recognition*, Vol. 31, No. 7, (1998), pp. 953-962.
[27] Yager N., Sowmya A., "Support vector machines for road extraction from remotely sensed images," *In Petkov, N.,Westenberg, M.A. (Eds.), Computer Analysis of Images and Patterns, Springer, Heidelberg*, (2003.), pp. 285-292.
[28] Yang C., Duraiswami R., Dementhon D., "Mean-shift analysis using quasi-newton methods," *In IEEE International Conference on Image Processing, Barcelona*, (2003).
[29] Yousif O., Ban Y.F., "Improving SAR-based urban change detection by combining MAP-MRF classifier and nonlocal means similarity weights," *Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, Vol. 7, No. 10, (2014), pp. 4288-4300.
[30] Zhu C.Q., Wang G.Y. and Ma Q.H., "Extracting roads based on morphological segmentation from RS image," *Journal of Surveying and Mapping*, Vol. 33, No. 4, (2004), pp. 347-351.
[31] Zhu D.M., Wen X. and Ling C.L., "Road extraction based on the algorithm of MRF and hybrid model of SVM and FCM," *In International Symposium on Image and Data Fusion, Tengchong*, (2011).