# Road Extraction from Remotely Sensed Data: A Review

Mohd Jawed Khan*, Pankaj Pratap Singh

Department of Computer Science and Engineering, Central Institute of Technology Kokrajhar, India

*Corresponding author

doi: https://doi.org/10.21467/proceedings.115.14

## ABSTRACT

Up-to-date road networks are crucial and challenging in computer vision tasks. Road extraction is yet important for vehicle navigation, urban-rural planning, disaster relief, traffic management, road monitoring and others. Road network maps facilitate a great number of applications in our everyday life. Therefore, a systematic review of deep learning approaches applied to remotely sensed imagery for road extraction is conducted in this paper. Four main types of deep learning approaches, namely, the GANs model, deconvolutional networks, FCNs, and patch-based CNNs models are presented in this paper. We also compare these various deep learning models applied to remotely sensed imagery to show their performances in extracting road parts from high-resolution remote sensed imagery. Later future research directions and research gaps are described.

**Keywords:** road extraction; machine learning; deep learning; remote sensing.

## 1 Introduction

The possession of detailed and reliable digital road network datasets can support a wide number of applications. Drone-based, airborne, spaceborne sensors using advanced Earth observation and remote sensing technologies. It obtained large amounts and different types of high-resolution images. Such images are extensively used in several applications, such as urban planning [1], disaster management [2], and emergency tasks [3]. Digital road network datasets are usually created by manual extraction initially. Although scientists have been trying to solve the problem of road detection for more than 30 years now [4], there hasn't been flawless software that can generalize the desired output under all different situations that occur in the built environment. It is a time consuming, expensive and labor-intensive procedure. However, manual road extraction is costly and slow, especially for rural areas which are weakly reachable and have limited access to GPS data unlike urban areas and highways [5], [6]. Instead, automatic road extraction from high resolution remote sensing (HR-RS) images have attracted much attention, since HR-RS images such as GanFen-2 (GF-2) have the advantages of wide imaging coverage, frequent revisit and sufficient spectral and spatial information of land covers [7], [8]. Most researchers regard road extraction from HR-RS images as a pixel-wise segmentation task which classifies each pixel into road or non-road according to geometric, photometric and texture features of roads (e.g. [9], [10], [11]). Thus, the performance of road extraction highly depends on the chosen semantic segmentation methods. In the last decade, the deep convolutional neural net- work (CNN) has been introduced to the road extraction due to its impressive performances in computer vision [12]–[14] level semantic features with consecutive convolutional operations. Recently, as a state-of-the-art machine learning technique, deep learning [15] made a major break-through in conventional computer vision tasks such as image classification, object detection, semantic segmentation and instance segmentation [16]. As a result, a large number of researchers started using deep



Proceedings DOI: 10.21467/proceedings.115; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-947843-5-7

learning techniques to solve remote sensing problems [17], including the road detection task [18]. Because of its supremacy of modeling complex nonlinear relationships between variables, deep learning surpassed conventional road detection algorithms [19]. However, the entire automation of the road extraction procedure is still not feasible. Extracted road networks using deep learning techniques frequently contain noise, artifacts, isolated road segments or miss information, making them inadequate for real-world applications .Mainly we are going in details of deep learning models: Patch-based CNNs, FCNs, Deconvolutional Nets, and GANs.

## 2 Deep Learning Models to RS Dataset

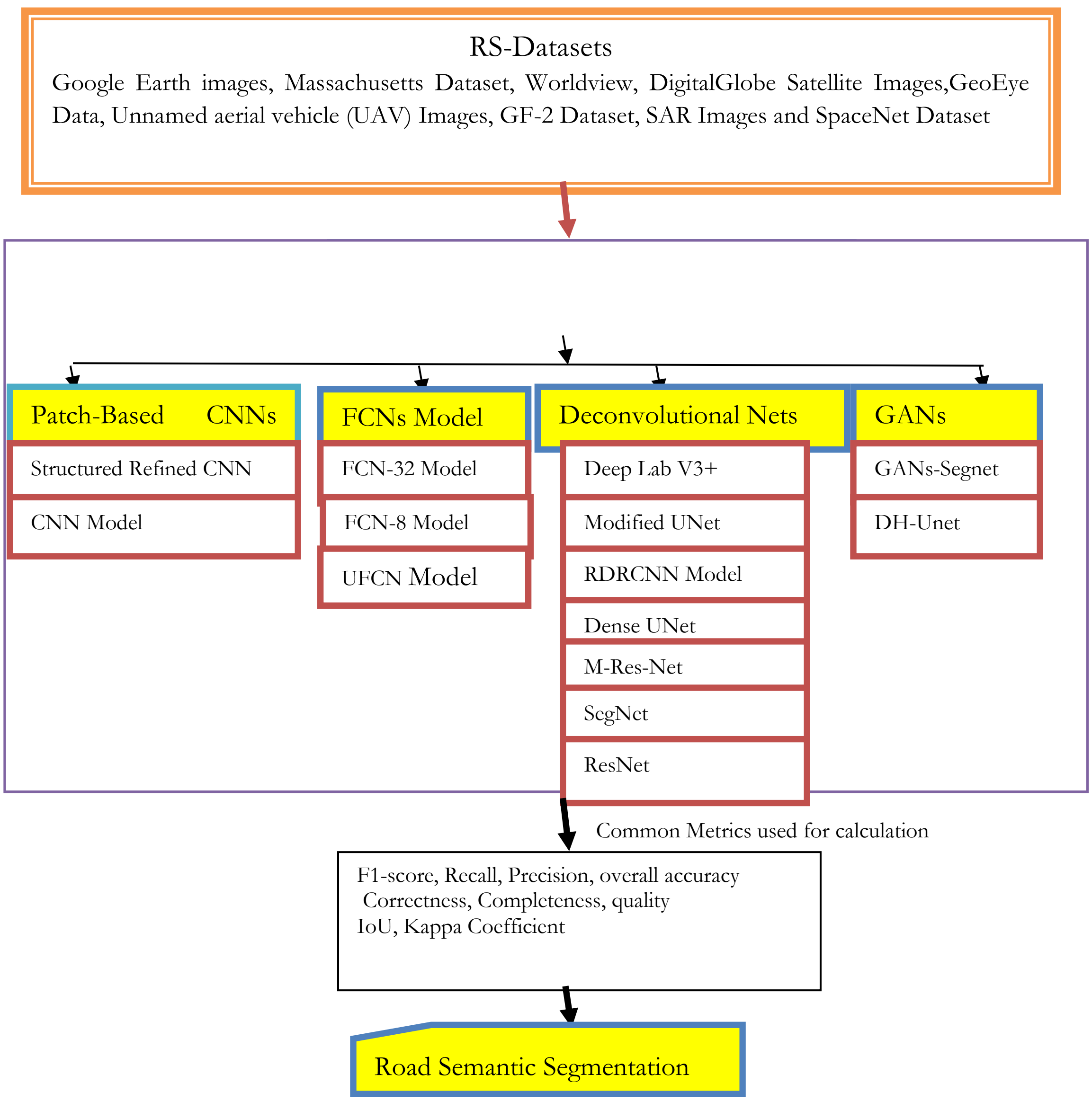


**Figure 1**. Semantic Segmentation of Road to RS Dataset using different Deep Learning Approaches.

### 2.1 Patch-Based CNN Model to Road Extraction

In the patch-based CNN model, the possibility of road dispensation is firstly predicted piece-by-piece with a particular stride and then the label map of the whole image is produced by assembling all of the label patches. Zhong et al. [20] provisionally implemented the newest CNN model to extract road and building objects from satellite imagery. The model fused low-level fine-grained features and high-level semantic meaning. In addition, further hyperparameters, such as the input image size, training epoch, and learning rate, were analyzed to specify the capability of the method in the context of high-resolution remote sensing images. The Massachusetts dataset, with a 1-m spatial resolution and 1500 × 1500 pixel size, containing 1711 images for the road and 151 images for the building datasets, was used for the evaluation. The Massachusetts dataset is related to the state of Massachusetts. The dataset covers over 2600 square kilometers with diverse rural, suburban, and urban areas [21]. With the integration of the pretrained FCN method with a novel four-stride pooling layer output to the last score layer, as well as fine-tuned with high-resolution spatial data, the extraction accuracy of the adjusted model was upgraded significantly to over 78%.

Wei et al. [22] applied aerial images for extracting road classes based on a road structure-refined CNN model, which provided road geometric information and spatial correlation. Furthermore, a novel road structure-based loss function was applied to cross-entropy loss to yield a weight map by using the minimum Euclidean distance of every pixel to the road section and to model the road geometric structure. The Massachusetts road dataset, including 1172 images randomly divided into 49, 14, and 1108 images for testing, validation, and training, respectively, was used to calculate the proposed technique. Efficiency measures, namely, F1 score, recall, precision, and accuracy, were calculated for comparison, which were 66.2%, 72.9%, 60.6%, and 92.4%, respectively.

### 2.2 Road Extraction Based on the FCNs Model

Compared to the CNN model that utilizes a dense layer to achieve a fixed-length feature vector and only accepts images with a fixed size, the FCNs model uses the interpolation layer after the final convolutional layer to upsample the feature map and restores the similar input size, as well as accepts input images of any size. In the FCNs, the final dense layers are replaced with convolutional layers and then output is a label map.

Varia et al. [23] applied a deep learning technique, namely, the FCN-32 for extracting road parts from extremely high-resolution UAV imagery. UAV-based imaging systems, which commonly use drones, can be used for the real-time assessment of several applications, monitoring tasks, and large-scale mapping, and are managed autonomously by onboard computers or remotely by human operators [24]. UAV-based remote sensing systems are used in various remote sensing applications, such as object recognition [25] and digital elevation model (DEM) generation [26]. Compared with traditional remotely sensed systems, UAVs have multiple advantages, including improved security, high speed, low cost, and high flexibility.

Proceedings DOI: 10.21467/proceedings.115
ISBN: 978-81-947843-5-7

Series: AIJR Proceedings
ISSN: 2582-3922

### 2.3 Deconvolutional Neural Networks (Dense Net) based Road Extraction

Deconvolutional networks struggle to extract hierarchical features from images that closely pertain to a number of deep learning methods from the machine learning community. These models comprise an encoder and decoder part, which a bottom-up mapping from the input image to the latent feature space is provided by the encoder part while the latent features are mapped back to the input image using the decoder part.

Panboonyuen et al. [27] presented a technique based on a modified deep encoder–decoder neural network to extract road objects from remote sensing imagery. To improve the suggested model, the authors enhanced certain phases of the suggested approach containing the incorporation of the exponential linear unit (ELU) function against the rectified linear unit function. In addition, the authors increased the number of training datasets by rotating images to eight different angles incrementally and used a landscape metrics (LM) method to eliminate false road parts and improve the general accuracy of the output. The designed model was tested on the Massachusetts dataset containing 49, 14, and 1108 images for testing, validation, and training, respectively. The most common metrics, namely, F1 score, recall, and precision, were also used for the performance evaluation, which gained 85.7%, 86.1%, and 85.4%, respectively. The results proved that the suggested approach yields satisfactory results and outperforms state-of-the-art approaches in road extraction from remote sensing imagery in terms of performance metrics.

Wang et al. [28] introduced a semiautomatic technique based on the finite state machine (FSM) and DNN, including two main steps, namely, training and tracking, for road extraction from high-resolution remote sensing imagery. In the training step, the model was trained to recognize the pattern of an input image. To generate training samples, a vector-guided labeling approach that elicited huge image-direction mates from available vector road maps and images was defined. In the tracking step, a fusion strategy was used to detect the size of a detection window, and the trained DNN was used to recognize extracted image patches. In general, the DNN was applied to the proposed method to determine a pattern from complicated scenes, and the FSM was used to control the behavior of trackers and translate identified patterns into states. The model was applied to two datasets, namely, aerial and Google Earth images, which were divided into 60%, 20%, and 20% for training, testing, and validation, respectively. Completeness, correctness, and quality percentage indices were used for the performance assessment, which were 75%, 70%, and 74%, respectively.

### 2.4 Road Extraction Based on the GANs Model

GANs comprises two main generative and discriminator models, in which the generative term tries to obtain the data dispensation and the discriminator part tries to determine the likelihood that a representation refers to training data instead of being created by a generative model [29]. Costea et al. [30] presented a new method named dual-hot generative adversarial networks (DH-GAN) to detect intersections and roads from UAV images at the higher semantic level of road graphs during the first step. Then, they applied a smoothing-based graph optimization method for pixelwise road segmenting and finding the road graph. They used F1 score, precision, and recall for evaluating the performance of the model, which were 86%, 89.84%, and 82.48% that proved the efficiency of the proposed model for road extraction, and also was able to minimize the memory costs. Shi, Liu, and Li [30] implemented the GANs model for attaining a smooth road segmentation map from Google Earth images with 550 images: 320 images were used for training, 100 images for validation, and 130 images for testing. They also used data augmentation procedures to increase the size of the dataset. An encoder–decoder SegNet model was used for the generative part to generate a high-resolution segmentation map. The accuracy that they achieved for recall, precision, and F1 score was 91.01%, 88.31%, and 89.63%, respectively, showing the superiority of the proposed model for road extraction.

## 3 Conclusion

Extracting road types using advanced remote sensing technologies can be economically and practically efficient. Numerous proposed methods for road extraction and road data updates using remote sensing images are described in this review. Therefore, the development of advanced machine learning methods, such as deep CNNs, for feature segmentation and extraction from remote sensing images has encouraged researchers. To apply such models to extract road networks from high spatial resolution remote sensing imagery, owing to the considerable efficiency of deep convolutional approaches in different applications. Following important outcomes are:

1. The capabilities of deep learning methods for road extraction are more effective than those of regular approaches.
2. When the complexity of images is high and various road types are present, the accuracy of the models is low. Therefore, mixing robust pre- and postprocessing techniques is recommended and useful to achieve satisfactory results.
3. The appropriateness of deep learning approaches for road extraction pertaining to different variables, such as architecture, data, and hyperparameters, is determined.
4. The low efficiency of the proposed methods in terms of data quality, training dataset, and model hyperparameters is presented.
5. Occlusions, such as shadows, cars, and buildings, are similar to road features, such as colors, reflectance, and patterns. Road extraction remains challenging owing to such issues.
6. In conclusion, several new methods related to road semantic segmentation are important and yet challenging, and research on different proposed techniques with cutting-edge technology is increasing.

## References

[1] Abdullahi, S.; Pradhan, B.; Jebur, M.N. GIS-based sustainable city compactness assessment using integration of MCDM, Bayes theorem and RADAR technology. Geocarto Int. 2015, 30, 365–387.

[2] Youssef, A.M.; Sefry, S.A.; Pradhan, B.; Alfadail, E.A.; Risk. Analysis on causes of flash flood in Jeddah city (Kingdom of Saudi Arabia) of 2009 and 2011 using multi-sensor remote sensing data and GIS. Geomat. Nat. Hazards 2016, 7, 1018–1042.

[3] Weng, Q. Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends.Remote Sens. Environ. 2012, 117, 34–49.

[4] Bajcsy, R. and Tavakoli, M. (1976). Computer recognition of roads from satellite pictures.IEEE Transactions on Systems, Man, and Cybernetics, 6:623–637.

Proceedings DOI: 10.21467/proceedings.115

ISBN: 978-81-947843-5-7



Series: AIJR Proceedings

ISSN: 2582-3922

[5] S. He, F. Bastani, S. Abbar, M. Alizadeh, H. Balakrishnan, S. Chawla, and S. Madden, ‘‘RoadRunner: Improving the precision of road network inference from GPS trajectories,’’ in Proc. 26th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst., Nov. 2018, pp. 3–12.

[6] E. Park, ‘‘Refining inferred road maps using GANs,’’ Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep., 2019.

[7] A Buslaev, S. Seferbekov, V. Iglovikov, and A. Shvets, ‘‘Fully convo- lutional network for automatic road extraction from satellite imagery,’’ in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2018, pp. 207–210.

[8] X.-Y. Tong, Q. Lu, G.-S. Xia, and L. Zhang, ‘‘Large-scale land cover classification in Gaofen-2 satellite imagery,’’ in Proc. IGARSS-IEEE Int. Geosci. Remote Sens. Symp., Jul. 2018, pp. 3599–3602.

[9] V. Mnih and G. E. Hinton, ‘‘Learning to detect roads in high-resolution aerial images,’’ in Proc. Eur. Conf. Comput. Vis. Berlin, Germany: Springer, 2010, pp. 210–223.

[10] X. Zhang, X. Han, C. Li, X. Tang, H. Zhou, and L. Jiao, ‘‘Aerial image road extraction based on an improved generative adversarial network,’’ Remote Sens., vol. 11, no. 8, p. 930, Apr. 2019.

[11] Z. Zhong, J. Li, W. Cui, and H. Jiang, ‘‘Fully convolutional networks for building and road extraction: Preliminary results,’’ in Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Jul. 2016, pp. 1591–1594.

[12] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 770–778.

[13] A Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘Imagenet classification with deep convolutional neural networks,’’ in Proc. Adv. Neural Inf. Pro- cess. Syst. (NIPS), 2012, pp. 1097–1105.

[14] M. Tan and Q. V. Le, ‘‘EfficientNet: Rethinking model scaling for convo- lutional neural networks,’’ 2019, arXiv:1905.11946. [Online]. Available: http://arxiv.org/abs/1905.11946

[15] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature, 521:436–44.

[16] Wei Liu and Dragomir Anguelov and Dumitru Erhan and Christian Szegedy and Scott Reed, and Cheng-Yang Fu, and Alexander C. Berg (2016). Ssd: Single shot multibox detector.

[17] Xu, Y., Xie, Z., Feng, Y., and Chen, Z. (2018). Road extraction from high-resolution remote sensing imagery using deep learning. Remote Sensing, 10(9).

[18] Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., and Eklund, P. (2016). A review of road extraction from remote sensing images.

[19] Batra, A., Singh, S., Pang, G., Basu, S., Jawahar, C., and Paluri, M. (2019). Improved road connectivity by joint learning of orientation and segmentation. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[20] Zhong, Z.; Li, J.; Cui, W.; Jiang, H. Fully convolutional networks for building and road extraction: Preliminary results. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 June 2016; pp. 1591–1594.

[21] He, K.; Zhang, X.; Ren, S.; Sun, J. Identity mappings in deep residual networks. In European Conference on Computer Vision; Springer: Berlin/Heidelberg, Germany, 2016; pp. 630–645.

[22] Wei, Y.; Wang, Z.; Xu, M. Road Structure Refined CNN for Road Extraction in Aerial Image. IEEE Geosci.Remote Sensing Lett. 2017, 14, 709–713.

[23] Varia, N.; Dokania, A.; Senthilnath, J. DeepExt: A Convolution Neural Network for Road Extraction using RGB images captured by UAV. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 18–21 November 2018; pp. 1890–1895.

[24] Abdollahi, A.; Pradhan, B.; Shukla, N. Extraction of road features from UAV images using a novel level set segmentation approach. Int. J. Urban Sci. 2019.

[25] Moranduzzo, T.; Melgani, F. Detecting cars in UAV images with a catalog-based approach. IEEE Trans. Geosci. Remote Sens. 2014, 52, 6356–6367.

[26] Yang, B.; Chen, C. Automatic registration of UAV-borne sequent images and LiDAR data. ISPRS J. Photogramm. Remote Sens. 2015, 101, 262–274.

[27] Panboonyuen, T.; Vateekul, P.; Jitkajornwanich, K.; Lawawirojwong, S. An enhanced deep convolutional encoder-decoder network for road segmentation on aerial imagery. In International Conference on Computing and Information Technology; Springer: Berlin/Heidelberg, Germany, 2017; pp. 191–201.

[28] Wang, J.; Song, J.; Chen, M.; Yang, Z. Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine. Int. J. Remote Sens. 2015, 36, 3144–3169.

[29] Shi, Q.; Liu, X.; Li, X. Road detection from remote sensing images by generative adversarial networks. IEEE Access 2018, 6, 25486–25494.

[30] Costea, D.; Marcu, A.; Slusanschi, E.; Leordeanu, M. Creating roadmaps in aerial images with generative adversarial networks and smoothing-based optimization. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 22–29 October 2017; pp. 2100–2109.