A new approach for road extraction using data augmentation and semantic segmentation

Kawther Ould Babaali[1,2], Ehlem Zigh[3], Mohamed Djebbouri[1], Oussama Chergui[4]
[1]Department of Electronics, Faculty of Electrical Engineering, University Djillali Liabes of Sidi Bel Abbes, Sidi Bel abbès, Algeria
[2]Department of Computer Science, M'Hamed Bougara University, Boumerdès, Algeria
[3]Department SPECIALITES, Laboratoire LaRATIC, Ecole Nationale Supérieure des Télécommunications et des TIC d'Oran, Oran, Algéria
[4]Department of Electronics, Laboratory Analysis of Signals and Systems, Mohamed Boudiaf University, M'sila, Algeria

Article Info

Article history:

Received Dec 28, 2021
Revised Aug 21, 2022
Accepted Sep 2, 2022

Keywords:

Data augmentation
Deep learning
Image classification
Remote sensing
Road extraction
Semantic segmentation

ABSTRACT

Accurate road extraction from remote sensing images is a challenging task. Several methods of extraction have been developed but the precision of extraction is still limited for the unpaved and small-width roads. This paper proposes an accurate road extraction approach called DAA-SSEG since it uses data augmentation architecture (DAA) and semantic segmentation model (SSEG). The proposed approach DAA-SSEG is based on a modified full convolutional neural network that overcomes the vanishing gradient and the training saturation issues. It recognizes roads at the pixel level. Furthermore, The DAA-SSEG approach uses a new plan of data augmentation based on geometric transformation and images refinement techniques. It allows getting a richer dataset thus better training and an accurate extraction. The experiment denotes that the proposed approach DAA-SSEG, that combine data augmentation architecture and semantic segmentation method, outperforms some state-of-the-art methods in terms of F-measures. The results demonstrate that it ensures accurate extraction of unpaved and small-width roads, in urban and rural areas. Moreover, the proposed approach distinguishes between roads and trails and can extract some roads not labeled beforehand.

Corresponding Author:

Kawther Ould Babaali
Department of Electronics, Faculty of Electrical Engineering, University of Djillali Liabès
Figuiers street BP89, 22000 Sidi Bel abbès, Algeria
Email: k.ouldbabaali@univ-boumerdes.dz

1. INTRODUCTION

Road extraction from high-resolution satellite and aerial images is one of the important assignments for instantaneous maps of emergency rescue systems, traffic monitoring in intelligent transportation, autonomous driving systems and real-time updating road networks [1], [2]. It is also used in many applications for planning urban areas and the footprint of buildings [3], [4]. Several researchers were interested in deep convolutional networks of machine learning because of their good abilities in automatic object extraction, and for their superiority to model complex non-linear relationships among variables. We can cite some proposed works that used convolutional neural network (CNN) models to detect roads, like the supervised multi-task learning model using recurrent convolution neural network U-Net. However, the model needs a huge amount of training samples to classify roads [5]. Another based-on CNN model for road extraction is The ResUnet. This model was developed to extract roads by using a deep residual U-net that shortened connections between the convolutional neural network's layers, to extend the U-net. However, ResUnet misses roads in the parking lot, considering them as a background [6].

Among several existing models, we can, also, cite the deep CNN based-on model that uses a cascade architecture with two levels to detect roads and then extract centerlines by using the domain area knowledge [7]. On the other hand, the model that uses classification method to detect road from satellite images with the help of CNN by using AlexNet with 128×128 resolutions in the Fourier domain, achieves also good results [8]. Further, another interesting model is the RoadTracer [9]. It can identify road by constructing a road network graph using an iterative graph construction method based on CNN. However, the model fails at road crossing positions and ignores high curvature and long straight roads.

This paper proposes a new approach for extracting road using a combination of data augmentation techniques and an improved semantic segmentation model. The proposed approach is called data augmentation architecture-semantic segmentation (DAA_SSEG).

Semantic segmentation is the process of assigning a label to every pixel in the image [10]. In contrast to the "classical" classification, where one label is assigned to the whole image. It is a precise-method well appreciated in medical image processing [11] and road extraction for autonomous driving [12].

Therefore, focusing on works that used semantic segmentation and deep learning for road extraction [13], we can cite some researches with good results, such as the deep CNN framework using conditional random fields (CRFs) and landscape metrics (LMs) [14] or the SegNet model that uses Exponential Linear Unit (ELU) as an activation function on a basic segmentation network [15]. SegNet demonstrates a promising detection performance on overall classes.

Moreover, another model, that has the ability to extract semantic maps of highways and road, is developed in [16]. It can also track the city's urban growth from remote sensing imagery. It is VGG-16 model with convolution kernels (3x3) and maximum pooling (2x2) across the model as a reference point for a fixed feature extractor. Prediction performance F1-scores was 0.76 on the Massachusetts Roads dataset.

Furthermore, one more interesting model named GL-Dense-U-Net was developed to extract global and local roads information from the remote sensing images [17]. The model uses U-Net and DenseNet as the feature extractor but it fails to extract unpaved road, due to the leak of the unpaved roads images in the tested dataset. The proposed approach is able to overcome almost all issues encountered in the state of the art such as the need for high number of input images, the false extraction of roads [18], the no extraction of parking roads and the confusion between roads and trails in [19].

2. MATERIAL AND METHODS

In this study, we propose an approach that associates a data augmentation architecture (DAA), suggested to use as a pre-processing to generate more data from input images, and an improved semantic segmentation based on a modified fully convolutional network (FCN) model (SSEG). The network uses a rectified linear unit (ReLu) activation and a binary cross entropy with a Logits Loss. As the relationship between the class to predict and the inputs (convoluted image segments) is 'non-linear', the performance of the network will improve if the fully connected layers (decision function) have non-linear activation functions. For that, we used a ReLU activation that overcomes the vanishing gradient problem, allowing models to learn faster and perform better. The binary cross entropy with a logits loss function ensures a numerical precision and a network stability. Thus, the semantic segmentation model (SSEG) of the proposed approach DAA-SSEG contains a less number of hidden layers in order to reduce the losses.

The DAA-SSEG approach includes two main parts: a data augmentation and a semantic segmentation proposed method. To show the efficiency and the accuracy of the DAA-SSEG approach we tested it on Massachusetts roads data set [20]. The data augmentation architecture is applied to the input images, the proposed model propagates the input images through the FCN, and then, it corrects the output mask using a probability score map thresholder with the removal of small objects and the filling of small holes to get finally extracted roads. The proposed road extraction approach DAA-SSEG in this paper, has been illustrated in Figure 1.

2.1. Tools and libraries

We have used the Tensorflow python framework [21] for our experimentation. Tensorflow is an open-source platform for artificial intelligence. The proposed model is trained according to the following configuration: learning rate of 0.001, with Adam optimizer (adaptive moment estimation) [22] to optimize the training process. To prevent the graphical card overloading, we have used a batch size of 16. In addition, we have used a single NVIDIA Tesla K80 GPU to train the network.

2.2. Dataset

To evaluate the proposed approach, we have used one of the most challenging available free dataset, the Massachusetts roads dataset [20]. Released by Toronto University, in Canada, it includes 1171 images with 1500x1500 pixels resolution. The whole dataset covers rural, urban and suburban regions with entire area over 2600 square kilometers. Each image covers a part of 2.25 square kilometers.

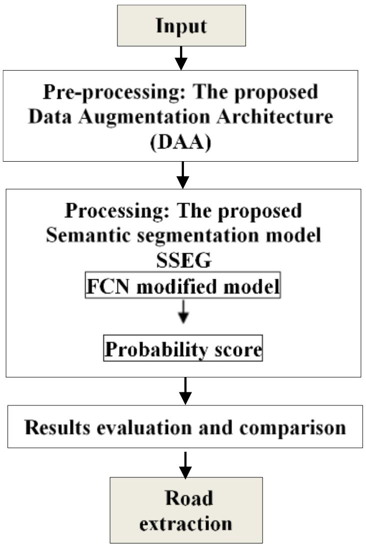


Figure 1. The proposed road extraction approach DAA-SSEG

2.3. Pre-processing

The proposed approach DAA-SSEG consists of pre-processing (data augmentation) and processing (semantic segmentation) parts. In pre-processing part, we have used a data augmentation technique, under the assumption that we can extract more information from the original dataset [23]-[25]. Data augmentation is used to increase the amount of input images by adding slightly modified copies of original images, or create new synthetic images from the original ones. The most popular data augmentation techniques are color modification, image refinement transformations and geometric distortions. Image refinement techniques are rotating in different degrees, reflecting, scaling with zoom (in and out), and cropping. Geometric distortions techniques are used to increase the number of samples for training, we can cite the histogram equalization and the enhancement of contrast and brightness with different levels [26].

To generate much more images from the Massachusetts dataset, which contains only 1171 images, we chose some image refinement and geometric transformation techniques as a data augmentation process. The dataset contains some damaged images, having white areas. To overcome this issue, the flawed images are selected and deleted completely, if necessary, or only the damaged part.

The proposed data augmentation architectures, denoted DAA, contains refinement transformations, like cropping, rotating, scaling (zoom in and zoom out) and geometric distortion like enhancing brightness. Several experiments were done to choose the appropriate refinement transformations and geometric distortion for optimal database inputs, and adapted image resolution. The proposed DAA for the pre-processing step has been illustrated in Figure 2.

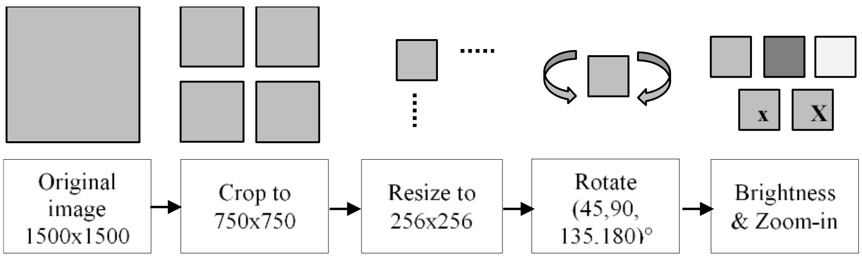


Figure 2. Proposed data augmentation architectures (DAA) for pre-processing

The first step of DAA is to crop each original image of 1500x1500 from the dataset into four images of 750x750. After that, the damaged images that contain white areas are removed partially or completely. Next step is to resize the remaining images into a resolution of 256x256 pixels in order to decrease the model training speed and keep images with well-described roads, especially roads with small width. Notice that resizing to smaller resolution makes small roads disappear.

Finally, geometric distortions and image refinement transformations are put on 200 images, to increase the number of training inputs by using scaling change (zoom-in), rotation in (45,90,135,180)° and brightness modification. Thus, a richer and more diverse dataset, with much more variations, is created. The total number of images gotten is 6172.

2.4. Model architecture

In this paper, we have proposed a model that is a fully connected convolutional network with a simplified structure layer to ensure a semantic segmentation. The model is called SSEG. Thus, each input pixel is associated with its class label. The network is inspired from U-net [27], developed to perform multiclass segmentation of any kind of object, primarily used for object in medical images [28], [29], and recently for aerial object [30] such as building, car or even roads. The input of the network is a multi-channel image and the output is a single channel map of the same dimension. We have trained SSEG with the pre-processed images created beforehand with the data augmentation architectures DAA.

The architecture of the proposed model is symmetric and uses an auto-encoder that learns data representations from input images in an unsupervised way. The structure of the proposed auto-encoder is composed of encoder, which learns the dense representation of input data to extract feature maps, and decoder, which decompresses it to reconstruct the input data in order to restore the feature maps resolution. The Architecture of the proposed FCN model used in the semantic segmentation method has been illustrated in Figure 3.

The encoder (downsampling) uses convolutional layers to learn different shapes of road in image. It uses downsampling part provided by the pooling layer. The model contains a pair of two or three convolution layers with a rectified linear unit (ReLu) activation, inspiring by [31] who used ReLu in deep learning. The purpose of applying the rectifier function (ReLu) is to increase the non-linearity in images. It is a simple calculation that returns the value provided as input directly, or the value 0 if the input is 0 or less.

Batch normalization and dropout layers follow the convolution and ReLu layers, to prevent overfitting and to ensure the independence of learning between layers. The batch normalization is used to accelerate the learning process of the model by decreasing the internal covariate shift.

To avoid overfitting, a dropout of keep probability of 0.3 is used. The dropout is a stochastic regularization technique that helps to reduce overfitting [32]. In addition, the padding used in the model is one (padding =1). The 'padding' is the amount of pixels added to an image when this last one is being processed by the kernel of the convolutional neural network (CNN).

The decoder (upsampling) uses a transposed convolution layer, which is a little bit different to a deconvolution layer. A transposed convolutional layer carries out a regular convolution but reverts its spatial transformation, to perform upsampling by 2 (stride=2) and filtering at the same time, with a kernel size of 4 and a padding of 1. The road extraction is ensured by a binary classification. At the end of the model, a (1x1) convolution layer is created to adapt the number of channels (exactly the third dimension) to the number of classes to segment or to combine the third dimension of the input feature maps to the two classes, road and background. As a sigmoid function, followed by a Binary Cross Entropy function, has the inconvenience of training saturation, we have overcome this issue, by using a binary cross entropy with logits loss that uses sigmoid internally, in order to train the network. For that, we have used the Binary Cross-entropy with logits loss (BCEWithLogitsLoss) referred in (3) which creates a criterion that measures the binary cross entropy (BCE), refered in (2) between the target and the output refered in (1). The BCEWithLogitsLoss referred in (3) combines the sigmoid layer and the BCELoss in one single class. It is more stable than using a plain Sigmoid followed by a BCELoss as, by combining the operations into one layer to ensure numeric precision and stability.

$$f(s_i) = \frac{1}{1+e^{-s_i}} \tag{1}$$

Where s_i is the input of sigmoid layer and $f(s_i)$ is the output of sigmoid layer.

$$BCE = -\sum_{i=1}^{C=2} t_i \log\big(f(s_i)\big) \tag{2}$$

$$BCEwithLogitsLoss = -t_1 \log\big(f(s_1)\big) - t_2 \log\big(f(s_2)\big) \tag{3}$$

Where s_i is the score i for class C_i , $f(s_i)$ is the output of sigmoid layer, t_i is the ground truth for i class, and i is the number of classes. For two classes C_1 and C_2 (Road and Background); t_1 $[0, 1]$ and s_1 are, respectively, the ground truth and the score for C_1. For the second class C_2, the ground truth $t_2=1-t_1$ and the score $s_2=1-s_1$.

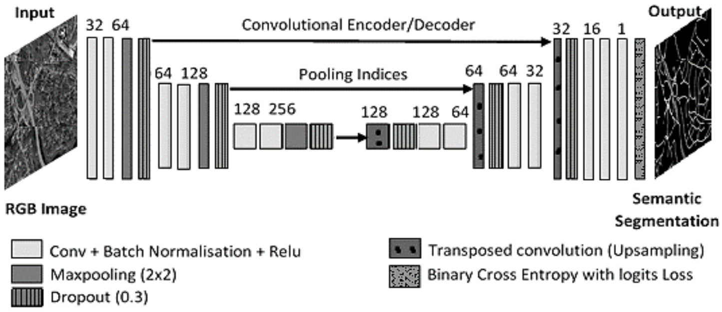


Figure 3. Architecture of the proposed SSEG model of the DAA-SSEG approach

3. RESULTS AND DISCUSSION

The quantitative and qualitative evaluations are important to show the efficiency and the precision of the DAA-SSEG approach for road extraction. In this section, we present at first the quantitative evaluation using different metrics then a comparison with existing baselines that used the same dataset. Finally, we end with a qualitative evaluation, which is a visual analysis of the resulting images.

3.1. Quantitative evaluation

The most common metrics for evaluating a binary classification method and semantic segmentation are precision, recall and F-measure. To show the importance of DAA used with the proposed model SSEG, Figure 4 illustrates the comparison between Precision-Recall curve of the DAA-SSEG approach and the SSEG model without using DAA. Notice that the inputs of this last one are only resized to 256x256. The F1-score of the model SSEG without using DAA is 0.70 because the number of input images is lowest (1088 images). Resizing input images from 1500x1500 to 256x256 pixels directly without using DAA make some information lost like roads with small width thus the extraction becomes much more difficult. The F1-score of the DAA-SSEG (training model SSEG using DAA as preprocessing) is 0.78.

Figure 4 shows that the DAA-SSEG presents a better F1-score thanks to the geometric distortions and the image refinement transformations used in the DAA. Indeed, DAA expands number of inputs by creating numerous images from the originals. The use of cropping, zooming, rotating in (0, 45, 90 ,135 ,180) and changing the brightness create a richer and diverse dataset with much more variations.

The F1-score of the DAA-SSEG is 0.78 and allows about 8% improvement compared to results of the training without the use of DAA. In terms of the computational cost, the DAA-SSEG required, about 9h. Notice that the use of ReLu and dropout layer in the SSEG model used in DAA-SSEG approach allows reducing overfitting and ensuring the independence of learning between layers.

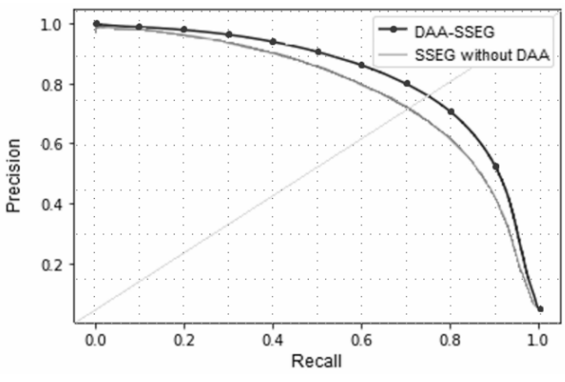


Figure 4. Precision-recall curve of the DAA-SSEG approach and the SSEG model without using DAA

3.2. Comparison with baselines

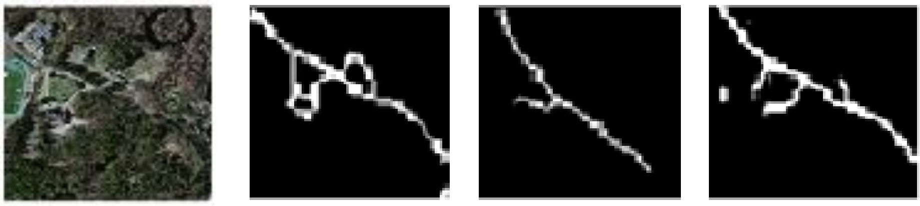
To demonstrate much more the efficiency and the accuracy of the proposed approach DAA-SSEG, Table 1 shows a comparison between the DAA-SSEG approach and four baselines (existing methods), counting Basic-model, FCN-no-skip, FCN-8s, [16] and basic SegNet run in [15]. Notice that the four baselines are tested on the Massachusetts dataset and the results are taken from the original paper [15] and [16]. Table 1 shows that the proposed approach DAA-SSEG outperforms the four baselines in term of F1-score, precision and recall.

Table 1. Comparison between results of baselines and the proposed approach DAA-SSEG tested on Massachusetts dataset [20]. (FCN: fully convolutional network)

Model	Precision	Recall	F1
Basic-Model [16]	0.657	0.657	0.657
FCN-no-skip [16]	0.742	0.742	0.742
FCN-8s [16]	0.762	0.762	0.762
SegNet [15]	0.773	0.765	0.768
DAA-SSEG	0.787	0.775	0.780

3.3. Qualitative evaluation

The qualitative evaluation shows the visual results after the processing. Figure 5 shows an example of roads extraction results using the proposed model SSEG without DAA and the DAA-SSEG approach. Figure 5(a) is the original image taken from the dataset, Figure 5(b) is its ground truth. The result of the extraction using the proposed model SSEG without DAA is shown in Figure 5(c) while the result of using SSEG with DAA (DAA-SSEG approach) is shown in Figure 5(d). we can see clearly that DAA-SSEG is much more efficient than training the proposed model SSEG witout DAA. It could extract more roads and is the most accurate.



(a) (b) (c) (d)

Figure 5. Example of visual results of roads extraction using SSEG without DAA and the DAA-SSEG approach, (a) original, (b) ground truth, (c) SSEG without DAA, and (d) DAA-SSEG

Figure 6 shows three examples of extraction of unpaved/unlabelled roads using DAA-SSEG and the distinction between error labels and real roads. It illustrates the efficiency and the accuracy of DAA-SSEG. Figure 6(a) is the original image of the first example taken from the Massachusetts dataset and Figure 6(b) represents its ground truth. The result of the roads extraction using DAA-SSEG is illustrated in Figure 6(c). It clearly shows that the proposed approach allows extracting more roads than the ground. It can extract roads with small widths such as roads leading to parking that are not labelled in the ground truth.

The same thing for the second example, Figure 6(d) is the original image and Figure 6(e) is its ground truth. Figure 6(f) is the result of the extraction using the DAA-SSEG approach. The last one denotes that even if, secondary unpaved roads and roads around parking are not labelled in the ground truth, DAA-SSEG allows a good extraction of those kinds of roads.

Like the previous example, Figure 6(g) is the original image, Figure 6(h) is its ground truth and Figure 6(i) is the result of the DAA-SSEG extraction. The most interesting in this example is the ability of the DAA-SSEG to distinguish between error labels and real roads. Indeed, in the ground truth, some labels contain a portion of regions considered as roads but are not really roads, as illustrated in the original image. The proposed DAA-SSEG approach offers the ability to correct this issue as demonstrated in the DAA-SSEG result. Consequently, DAA-SSEG is able to extract the paved and unpaved roads.

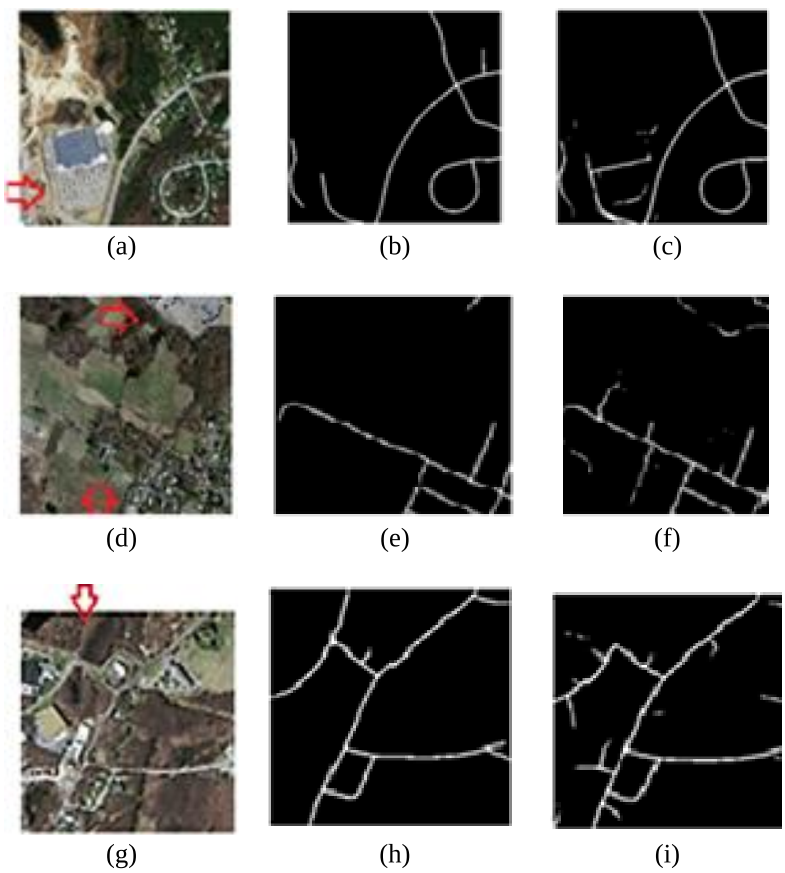


Figure 6. Examples of extraction of unpaved/unlabeled roads using DAA-SSEG and the distinction between error labels and real roads, (a) Original, (b) Ground truth, (c) DAA-SSEG, (d) Original, (e) Ground truth, (f) DAA-SSEG, (g) Original, (h) Ground truth, and (i) DAA-SSEG

4. CONCLUSION

This paper proposed a data augmentation and semantic segmentation approach DAA-SSEG for accurate road extraction from remote sensing images. It deals with the complex structures of roads that have different types and widths. The proposed approach used an optimized semantic segmentation based deep learning model with fewer layers than a classical FCN model to ensure recognition at pixel level. It contains an encoder/decoder with a Binary Cross Entropy with a Logits Loss function, in order to avoid a training saturation. The proposed semantic segmentation model ensures efficient and accurate road extraction, by building a deep pipeline that takes remote sensing images as input and produces semantically segmented maps. To show the importance and the contribution of the pre-processing step, the second contribution of this paper is the study of the impact of data augmentation technique on the result of semantic segmentation based deep learning proposed method. The quantitative and qualitative evaluations, demonstrate that the proposed approach DAA-SSEG outperforms some state of the art of methods. Indeed, using the binary cross entropy with logits loss (BCEWithLogitsLoss) function in the SSEG model of the DAA-SSEG approach allows a good extraction of roads. Moreover, using the proposed DAA as preprocessing before training SSEG model allows also getting a richer dataset thus a better training and an accurate extraction of roads with small width, secondary, falsely labelled and unpaved roads. Furthermore, the proposed DAA-SSEG approach distinguishes between roads and trails and can extract some roads not labeled beforehand. This demonstrates the crucial interest for updating some ground truth datasets. For future work, more choices of other neural network architectures or activation functions can be investigated and compared to improve much more the results.

REFERENCES

[1] J. Wang, J. Song, M. Chen, and Z. Yang, “Road network extraction: a neural-dynamic framework based on deep learning and a finite state machine,” *International Journal of Remote Sensing*, vol. 36, no. 12, pp. 3144–3169, Jun. 2015, doi: 10.1080/01431161.2015.1054049.

[2] J. Zhang, L. Chen, C. Wang, L. Zhuo, Q. Tian, and X. Liang, “Road recognition from remote sensing imagery using incremental learning,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 11, pp. 2993–3005, Nov. 2017, doi: 10.1109/TITS.2017.2665658.
[3] 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, pp. 3359–3372, Jun. 2014, doi: 10.1109/TGRS.2013.2272593.
[4] Y. Xu, L. Wu, Z. Xie, and Z. Chen, “Building extraction in very high resolution remote sensing imagery using deep learning and guided filters,” *Remote Sensing*, vol. 10, no. 1, p. 144, Jan. 2018, doi: 10.3390/rs10010144.
[5] X. Yang, X. Li, Y. Ye, R. Y. K. Lau, X. Zhang, and X. Huang, “Road detection and centerline extraction via deep recurrent convolutional neural network U-Net,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 7209–7220, Sep. 2019, doi: 10.1109/TGRS.2019.2912301.
[6] Z. Zhang, Q. Liu, and Y. Wang, “Road extraction by deep residual U-Net,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749–753, May 2018, doi: 10.1109/LGRS.2018.2802944.
[7] G. Cheng, Y. Wang, S. Xu, H. Wang, S. Xiang, and C. Pan, “Automatic road detection and centerline extraction via cascaded end-to-end convolutional neural network,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 3322–3337, Jun. 2017, doi: 10.1109/TGRS.2017.2669341.
[8] J. Hormese and C. Saravanan, “A convolutional neural network approach to road classification from satellite images,” *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 23, pp. 7917–7927, 2018.
[9] F. Bastani *et al.*, “RoadTracer: Automatic extraction of road networks from aerial images,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 4720–4728, doi: 10.1109/CVPR.2018.00496.
[10] M. Thoma, “A survey of semantic segmentation,” *arXiv:1602.06541v2*, 2016, [Online]. Available: http://arxiv.org/abs/1602.06541.
[11] I. Rizwan, I Haque, and J. Neubert, “Deep learning approaches to biomedical image segmentation,” *Informatics in Medicine Unlocked*, vol. 18, p. 100297, 2020, doi: 10.1016/j.imu.2020.100297.
[12] M. Treml *et al.*, “Speeding up semantic segmentation for autonomous driving,” *NIPS 2016 workshop MLITS*, no. Nips, pp. 1–7, 2016, [Online]. Available: https://openreview.net/pdf?id=S1uHiFyyg%0Ahttps://openreview.net/forum?id=S1uHiFyyg.
[13] F. Lateef and Y. Ruichek, “Survey on semantic segmentation using deep learning techniques,” *Neurocomputing*, vol. 338, pp. 321–348, Apr. 2019, doi: 10.1016/j.neucom.2019.02.003.
[14] 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 Sensing*, vol. 9, no. 7, p. 680, Jul. 2017, doi: 10.3390/rs9070680.
[15] V. Badrinarayanan, A. Handa, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling,” *arXiv:1505.07293*, 2015, [Online]. Available: http://arxiv.org/abs/1505.07293.
[16] S. Muruganandham, “Semantic segmentation of satellite images using deep learning,” Theses, Czech Technical University in Prague & Luleå University of Technology, 2016.
[17] Y. Xu, Z. Xie, Y. Feng, and Z. Chen, “Road extraction from high-resolution remote sensing imagery using deep learning,” *Remote Sensing*, vol. 10, no. 9, p. 1461, Sep. 2018, doi: 10.3390/rs10091461.
[18] M. Li, A. Stein, W. Bijker, and Q. Zhan, “Region-based urban road extraction from VHR satellite images using Binary Partition Tree,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 44, pp. 217–225, Feb. 2016, doi: 10.1016/j.jag.2015.09.005.
[19] V. Mnih and G. E. Hinton, “Learning to detect roads in high-resolution aerial images,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6316 LNCS, no. PART 6, 2010, pp. 210–223.
[20] V. Mnih, “Road and Building Projects,” Theses, University of Toronto, 2013.
[21] J. Paul, “Python machine learning: machine learning and deep learning with python,” *Towards Data Science*, 2019. https://towardsdatascience.com/road-segmentation-727fb41c51af (accessed Jun. 01, 2020).
[22] J. L. Ba and D. P. Kingma, “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015, pp. 1–15.
[23] M. D Bloice, C. Stocker, and A. Holzinger, “Augmentor: An image augmentation library for machine learning,” *The Journal of Open Source Software*, vol. 2, no. 19, p. 432, Nov. 2017, doi: 10.21105/joss.00432.
[24] H. Inoue, “Data augmentation by pairing samples for images classification,” 2018, [Online]. Available: http://arxiv.org/abs/1801.02929.
[25] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
[26] A. Mikolajczyk and M. Grochowski, “Data augmentation for improving deep learning in image classification problem,” in *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, May 2018, pp. 117–122, doi: 10.1109/IIPHDW.2018.8388338.
[27] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: convolutional networks for biomedical image segmentation,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9351, 2015, pp. 234–241.
[28] N. R. Shenoy and A. Jatti, “Ultrasound image segmentation through deep learning based improvised U-net,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, pp. 1424–1434, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1424-1434.
[29] X. X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, “U-Net-based medical image segmentation,” *Journal of Healthcare Engineering*, vol. 2022, pp. 1–16, Apr. 2022, doi: 10.1155/2022/4189781.
[30] A. F. Agarap, “Deep learning using rectified linear units (ReLU),” *arXiv:1803.08375v2*, 2018, [Online]. Available: http://arxiv.org/abs/1803.08375.
[31] B. Jabir and N. Falih, “Dropout, a basic and effective regularization method for a deep learning model: A case study,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 2, pp. 1009–1016, Nov. 2021, doi: 10.11591/ijeecs.v24.i2.pp1009-1016.

BIOGRAPHIES OF AUTHORS

Kawther Ould babaali is a Doctorate student in signal and image processing, at the electronics department of university Djillali Liabes of Sidi Bel Abbes, Algeria and works on an Image Processing and Hyperspectral classification project at the same university. She also works as teacher at department of computer science at University M'hamed Bougara of Boumerdes, Algeria. She gets her Magister degree in 2012 at Ecole Nationale Superieure des Télécommunications et des TIC d'Oran (ex INTTIC), Algeria. Her research interests related to remote sensing, artificial intelligence techniques and Algorithms. She can be contacted at email: k.ouldbabaali@univ-boumerdes.dz.

Ehlem Zigh is Associate Professor at Ecole Nationale Superieure des Télécommunications et des TIC ENSTTIC d'Oran (ex INTTIC), Algeria. She received her Habilitation degree in Electronic field from Djilalli liabess University Belabes, Algeria in 2017. Her Doctorate degree from the University of Sciences and Technologies Mohammed Boudiaf of Oran, Algeria in 2014. She is a Head of a research group at LaRATIC Laboratory at ENSTTIC. She has around 20 national and international communications, eleven international publications and she published a book chapter in a Handbook of Research on Artificial Intelligence Techniques and Algorithms. Malaysia. She is a reviewer at The International Journal of Energy Optimization and Engineering, IGI Global, European journal of remote sensing and Interactive Learning Environment journal. She can be contacted at email: ezigh@inttic.dz.

Mohamed Djebbouri is a Professor at University of Sidi-Bel-Abbes, Department of Electronics from 1991, his domain of interest is digital signal processing, imagery, embedded systems, mobile satellite communications, and GNSS. He can be contacted at email: mohamed.djebbouri@univ-sba.dz.

Oussama Chergui is a Doctorate student at Université de M'sila and researcher at System and signal processing laboratory at Université de M'sila. He obtained his Master degree at University Yahia Fares of Médéa, Algeria. his domain of interest is digital signal processing, artificial intelligence. He can be contacted at email: oussama.chergui@univ-msila.dz.