International Journal of Image and Graphics
Vol. 22, No. 4 (2023) 2350058 (25 pages)

DOI: 10.1142/S0219467823500584

World Scientific
www.worldscientific.com

Deep Learning-Enabled Road Segmentation and Edge-Centerline Extraction from High-Resolution Remote Sensing Images

Miral Jerambhai Patel* and Ashish M. Kothari

Research Scholar, Atmiya University
Assistant Professor, Government Engineering College
Rajkot, Gujarat, India

Department of Electronics & Communication Engineering
Atmiya University, Rajkot, Gujarat, India
**mjpatelgec@gmail.com*

Received 1 February 2022
Accepted 22 June 2022
Published

Nowadays, precise and up-to-date maps of road are of great significance in an extensive series of applications. However, it automatically extracts the road surfaces from high-resolution remote sensed images which will remain as a demanding issue owing to the occlusion of buildings, trees, and intricate backgrounds. In order to address these issues, a robust Gradient Descent Sea Lion Optimization-based U-Net (GDSLO-based U-Net) is developed in this research work for road outward extraction from High Resolution (HR) sensing images. The developed GDSLO algorithm is newly devised by the incorporation of Stochastic Gradient Descent (SGD) and Sea Lion Optimization Algorithm (SLnO) algorithm. Input image is pre-processed and U-Net is employed in road segmentation phase for extracting the road surfaces. Meanwhile, training data of U-Net has to be done by using the GDSLO optimization algorithm. Once road segmentation is done, road edge detection and road centerline detection is performed using Fully Convolutional Network (FCN). However, the developed GDSLO-based U-Net method achieved superior performance by containing the estimation criteria, including precision, recall, and F1-measure through highest rate of 0.887, 0.930, and 0.809, respectively.

Keywords: Road surface segmentation; road edge detection; road centerline detection; sea lion optimization algorithm; and stochastic gradient descent.

1. Introduction

In recent times, information regarding the roads has become vital as they support numerous applications, like urban planning, vehicle routing system, change recognition and hence road surface extraction from images with high resolution is highly significant. However, the roads are modeled as a group of intersections and connection among these intersections.[2] At present, road surface segmentation in remote

*Corresponding author.

sensing images has turned out to be the critical job in various urban applications, like urban planning, road monitoring, and traffic management.[7] In remote sensing, automatic extraction of roads has become an open and active research issue.[35,36] It is a fundamental pre-processing stage for applications such as geographic information update, image filling.[32,33] On the other hand, it is a monotonous and time-consuming procedure for labeling the road areas manually. With the support of artificial intelligence, image processing, and machine learning, automatic extraction of roads from remotely sensed images is an cost-effective and successful mode to obtain the road information.[3,16,21,22] With an enhancement of Convolution Neural Networks (CNNs), recognition of roads from remote sensing images tends to be an effectual procedure.[7] In general, the road extraction process comprises two subdivisions, namely, road edge detection and extraction of road midway. The former aims at extracting every road pixels out, whilst the latter is to label the pixels of road midway, which can give instructions for the vehicle routing and other alternative applications.[3]

Remote sensing images obtained by space borne or airborne sensors are the key source for object recognition, earth surface monitoring and environmental inspection. In general, the remote sensing images with HR have become highly significant for disaster monitoring, biological study, geographic information system (GIS) application, and land management. The key concern is how to detect the objects of interest in HR remote sensing images accurately and more quickly. The basic processing of HR remote sensing images is to extract the required details for objects recognition and classification based on the shape characters, spatial relations, and spectral features.[9,11,29] In general, trees, buildings, and cars all along the roads form shadows or can cause obstructions that create heterogeneous areas on roads. Therefore, road detection remains as a difficult process.[7] Road surface segmentation generally focuses on creating a binary mask map where every pixel is labeled as either non-road or road.[6] In the urban regions, surfaces of roads are normally small altitude flat jumps caused by curb stones on the restrictions of road. Moreover, the elevations of road boundaries change more quickly to a certain extent than road surface points. It is important to divide roads from other points for achieving accurate segmentation of road surfaces as the scalar field gradients reflect the rate of scalar change.[11,12] There are various methods, like Dual Attention Network (DANet),[37] Criss-Cross Network (CCNet),[38] Attention Complementary Network (ACNet),[39] SEgmentation TRansformer (SETR),[40] SegFormer,[41] Trans4Trans,[42] and Aerial-Panoramic Annular Scene Segmentation (Aerial-PASS)[43] are used in the effective segmentation. Furthermore, Road edge detection plays a major role for detecting the spot of obstacles, road direction, speed and size of impediment in the road surface.[13–15] Consequently, road midway extraction concentrates on the connectivity and topology of road networks, which is normally administered by thinning or line tracking from road mask maps.[6,17]

According to the extracted road outcomes, most of the researches have concentrated on road region extraction and road centerline extraction.[7,18–20] Several

techniques are introduced for road detection, but mainly they are using pixel-level labeling. Because of the occlusions and noises under trees and cars in the remote sensing images with very high-resolution, the existing approaches often generate diverse heterogeneous outcomes.[3] Moreover, deep learning methods have been gradually used to extract the information from satellite images with high-resolution due to their generalization capability and superior performance.[10,34] Deep learning techniques[30,31] have been used for object detection, semantic segmentation, and image categorization, and various other computer vision tasks provide a promising way for road exterior extraction from remote sensing images. In past decades, numerous techniques for road detection have been highly developed from diverse viewpoints that are normally partitioned into two different sets, such as road superficial segmentation and extract the road midway. Furthermore, conventional information and machine learning approaches, namely, support vector machine (SVM), artificial neural network (ANN), and maximum likelihood (ML) have been extensively established in the extraction of road surface. Recently, CNN has turned out to be outstanding in semantic segmentation of visual images. Motivated by the thriving applications related to deep learning schemes in semantic segmentation, a few studies have established CNNs, particularly FCNs into road surface extraction.[6]

The major motivation of this work is to establish a deep learning-enabled road segmentation and extraction of edge-midway from HR remote sensing images. At first, input is presented to pre-processing step to remove the noises from the input. With pre-processed image, road segmentation is performed for extracting road surfaces using U-Net that is accomplished by the established GDSLO algorithm. Once road segmentation is done, road edge detection and road centerline extraction is done using FCN based on the pre-processed input image and road segmented output. Moreover, the proposed GDSLO is devised with the amalgamation of SGD and SLnO.

The primary objective of this research work is demonstrated as follows:

- **Proposed GDSLO-based U-Net:** An efficient and robust optimization algorithm, named GDSLO is introduced for road segmentation and also for training U-Net for achieving better performance in extracting the road surfaces. Moreover, the developed GDSLO is the hybridization of SLnO and SGD.

The organization of the paper is as follows: The numerous prevailing deep learning approaches for the segmentation and extraction of the road surface are reviewed in Secs. 2 and 3 portrays the explanation of the proposed model. Section 4 portrays results and discussion. Finally, Sec. 5 is the conclusion of the paper.

2. Motivation

The several existing deep learning techniques based on road surface segmentation and extraction as well as its merits and drawbacks are described in this segment that motivate researchers to design the proposed GDSLO-based U-Net.

2.1. *Literature review*

Various existing deep learning approaches based on road exterior segmentation and extraction are discussed in this section. Liu *et al.*[1] designed a Deep CNN model for predicting road surfaces, edges, and midway. This approach resolved the occlusions and shadows along the road regions. However, this method failed to use refining techniques, like guided filtering and Conditional random fields (CRFs) to increase the performance. Liu *et al.*[2] introduced a CNN for extracting road midway. Here, multi-scale Gabor filters and multi-directional non-maximum suppression was accomplished for extracting road midway accurately. For extracting the road midway, this process was more effective, but this method is unsuccessful to diminish the processing complexity. Cheng *et al.*[3] devised a cascaded end-to-end convolutional neural network (CasNet) for extracting surfaces and centerlines of the roads. Here, CNN was utilized to extract consistent road area outcomes. This method achieved extraordinarily smooth and consistent road detection results. However, CasNet failed to cut down the computational complexities. Senchuri *et al.*[4] developed a Multisensory hyperspectral and LiDAR data fusion (HL-Fusion). This method achieved robust performance with higher processing speed and less over fitting issues. Meanwhile, this approach does not consider the novel deep learning algorithms for achieving effective outcomes.

Chen and Chen[5] presented a Road and road Boundary detection Network (RBNet) for detecting road and boundaries of road. This method estimated the probability of the image pixels related to road and edges of the road. In a single processing stage, boundaries and areas of the road are detected by using the RBnet. But this technique failed to improve the computational speed. Wei *et al.*[6] designed an FCN model for extracting the surfaces and centerlines of road. This model accurately detected the road networks and hence achieved improved extraction accuracy. Besides, this model failed to identify road and midway with lesser number of training samples. Shamsolmoali *et al.*[7] devised a spatial pyramid (SP) network for road segmentation. Here, a feature pyramid network was established to take out the significant characteristics from the network layers for representing different object scales. This approach achieved effective performance in terms of effectiveness and accuracy. It caused some displacement errors corresponding to the ground truth. Wei and Ji[8] designed a Scribble-based weakly supervised road surface extraction method (ScRoadExtractor) for extracting the surfaces of road. This method attained better performance without the need for alternating optimization algorithms. However, this method was not utilized for other types of segmentation tasks.

2.2. *Challenges*

The numerous issues met through the deep learning approaches on the basis of road surface segmentation and extraction are presented below as follows:

- In Ref. 2, CNN was developed for extracting the centerlines of roads from high-resolution aerial images precisely. However, some of the midways extracted by the method were not single-pixel wide.

- The CasNet developed in Ref. 3 achieved robustness against the occlusions of trees, but the major challenge lies in detecting the roads accurately by incorporating more high-level semantic information to extract large and continuous area of occlusions.
- The difficult backgrounds of remote sensing images made the road surface extraction practice to suffer, namely, overlapping of viaducts and occlusion of buildings and trees.[6]
- The ASPN developed in Ref. 7 attained better performance using domain adaptation and adversarial segmentation. However, it failed to construct a more effective approach for improving the accuracy of segmentation not only on road but also on the segmentation procedures while minimizing total cost of computation.
- The construction of graphs in the road label propagation point can be further expressed as graph representation learning, which can entrench the topological facts of road networks into the learning-enabled graph models.[8]

3. Proposed Deep Learning-Enabled Approach for Road Segmentation and Edge-Centerline Extraction

The major role of the research is to aim and develop a method, named as GDSLO-based U-Net for road superficial extraction from HR remote sensing images. This approach comprises of various phases like pre-processing, division of road surface, detection of road edge, and extraction of road centerline. At first, the input image is passed to the pre-processing phase where the image is pre-processed to remove the noises and external artifacts by employing T2FCS filter.[24] After pre-processing, the pre-processed image is presented to the road segmentation phase where the road surfaces are effectively extracted using U-Net.[27] However, the training of U-Net is achieved by means of established optimization algorithm, called GDSLO. The proposed GDSLO is designed by integrating SGD with SLnO.[23] Once road segmentation is done, road edge detection is carried out using FCN[28] based on the pre-processed result and segmented output. Finally, the road centerline is effectively extracted by considering the pre-processed input image and road segmented result. Figure 1 demonstrates the diagrammatic illustration of the proposed system.

3.1. *Acquisition of input*

Input is gathered from the database and the set of data D is presumed with α amount of input, which can be represented as

$$D = \{S_1, S_2, \ldots, S_z, \ldots, S_\alpha\}, \tag{1}$$

where D symbolizes dataset, whole number of images is specified as α, and S_z represents the input image at zth index, which is presented as an input to pre-processing stage.

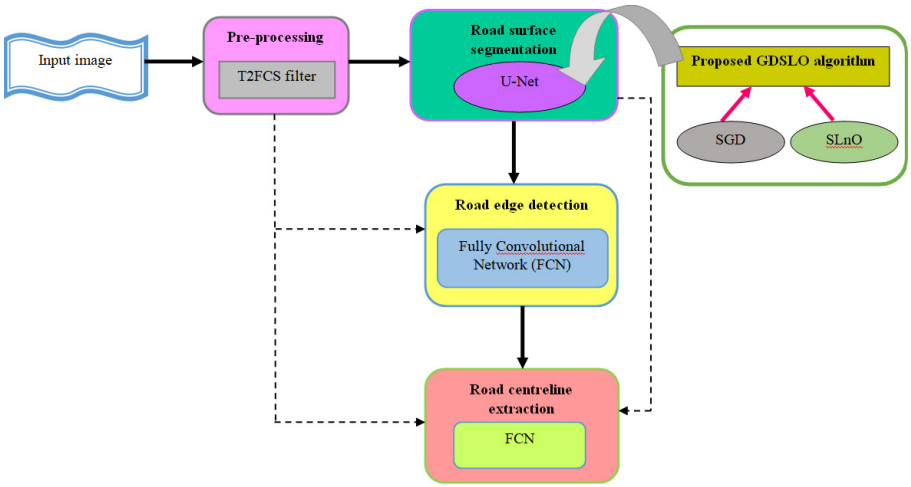


Fig. 1. Schematic view of the established deep learning-enabled method for road segmentation and edge-midway extraction.

3.2. *Pre-processing using T2FCS filter*

The aim of the pre-processing is to remove falsifications or noises from the input. Here, T2FCS filter[24] is employed for pre-processing the input image. T2FCS filter is designed by integrating Cuckoo Search (CS) optimization in type II fuzzy system in which each associate is denoted as 0 and 1 based on fuzzy system. Here, type 2 fuzzy groups are altering the image without damaging the image pixels to make noise free. Moreover, large amount of gradation of uncertainty is controlled by means of type 2 fuzzy system. Consequently, T2FCS filter helps to detect noises from distorted input. P_i denotes pre-processed image result, which is specified to the road segmentation.

3.3. *Road segmentation using GDSLO-based U-Net*

In this phase, road segmentation is done for extracting the road surfaces. Segmentation is the method for separating the pre-processed output into various sections that is image pixels or objects of roads. Here, the road segmentation process is carried out by the U-Net model.[27] However, the training practice of U-Net is done using the proposed GDSLO algorithm, which is the hybridization of SGD[44] and SLnO.[23]

3.3.1. *Architecture of U-Net*

The U-Net architecture[27] comprises of two different paths, such as contracting path on the left side and expansive path on the right side. Here, the contracting path pursues the convolution network typical model. The appropriate data is gathered from the pre-processed output P_i through the contracting path, whereas the apt region to be segregated is localized with the expansive path. From the contracting

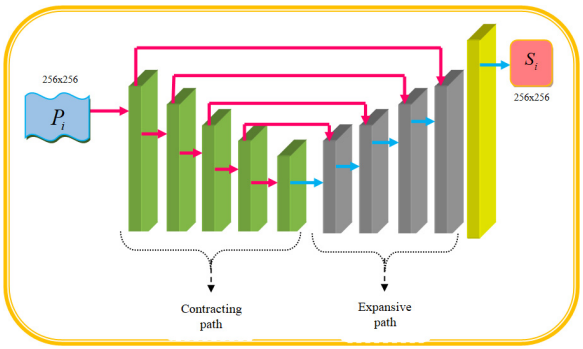


Fig. 2. Architecture of U-Net.

path, the features with the higher-level values of pixels are generated which is then incorporated with the characteristic maps during the process of upsampling to restore the images. The entirely linked layers considerably minimized the required constraints to train the neural network with the minimum amount of data. Every stage of expansive path contains feature map upsampling, which is followed by up-convolution that halves the overall feature channels, thereby obtaining segmented output by transforming the feature map channels. Here, the stimulation function is accomplished in the Rectified Linear Unit (ReLU). Moreover, in each convolution, cropping procedure is significant because of the loss of border pixels. As a result, the segmented image result is signified as S_i. The architectural representation of U-Net model is shown in Fig. 2.

3.3.2. *Training practice of U-Net using developed GDSLO algorithm*

This section elucidates the training process of U-Net[27] using proposed optimization algorithm, called GDSLO. Here, the developed GDSLO is newly designed by the hybridization of SGD and SLnO.[23] The SLnO[23] algorithm is inspired by the hunting performance of sea lions. Furthermore, it is driven by whiskers of sea lions' to recognize target. The SLnO process attains improved generalization capability and evades premature convergence such that the optimal solution can be effectively identified. Meanwhile, SGD[44] is an iterative approach used for optimizing the objective function with proper smoothness properties. Moreover, the integration of SGD and SLnO enhances the overall performance of the system and attains optimum resolution by reducing the computational issues. The algorithmic stages of proposed GDSLO algorithm are presented as follows:

(1) *Initialization*

Solution Initialization of P can be expressed as

$$P = \{P_1, P_2, \ldots, P_m, \ldots, P_z; 1 \leq m \leq z\}, \tag{2}$$

where z denotes complete solutions, and P_m signifies the mth result.

(2) *Fitness evaluation*

Fitness function estimates the optimum solution including the fitness function with the minimum error which is considered as finest solution, and fitness equation can be formulated as

$$F = \frac{1}{\ell} \sum_{\delta=1}^{\ell} [T - O_r^*]^2, \tag{3}$$

where F denotes fitness measure, O_r^* represents output generated by U-Net, and T implies target outcome.

(3) *Detecting and tracking phase*

According to SLnO[23] algorithm, the whiskers of sea lions are used for predicting the shape, size, and location of prey. SLnO considers the target as current optimum solution or closer to optimum solution. In addition, sea lion travels along with the direction of target prey and formulated as

$$\mathbf{P}(g+1) = \mathbf{S}(g) - \mathbf{L} \cdot \mathbf{Q}, \tag{4}$$

$$\mathbf{P}(g+1) = \mathbf{S}(g) - |2\mathbf{M} \cdot \mathbf{S}(g) - \mathbf{P}(g)| \cdot \mathbf{Q}. \tag{5}$$

Consider $\mathbf{M} > 0.5$

$$\mathbf{P}(g+1) = \mathbf{S}(g) - (2\mathbf{M} \cdot \mathbf{S}(g) - \mathbf{P}(g)) \cdot \mathbf{Q}, \tag{6}$$

$$\mathbf{P}(g+1) = \mathbf{S}(g)(1 - 2\mathbf{M} \cdot \mathbf{Q}) + \mathbf{P}(g) \cdot \mathbf{Q}. \tag{7}$$

By incorporating SLnO with SGD, the optimum solution can be achieved by minimizing the optimization issues. The standard equation of SGD is given as

$$\mathbf{P}(g+1) = \mathbf{P}(g) - \upsilon_h \nabla R(P(g)), \tag{8}$$

$$\mathbf{P}(g) = \mathbf{P}(g+1) + \upsilon_h \nabla R(P(g)), \tag{9}$$

where the learning rate is denoted as υ_h, $\nabla R(P(g))$ indicates gradient of the position vector loss function.

By substituting (9) in (7), the equation becomes

$$\mathbf{P}(g+1) = \mathbf{S}(g)(1 - 2\mathbf{M} \cdot \mathbf{Q}) + (\mathbf{P}(g+1) + \upsilon_h \nabla R(P(g))) \cdot \mathbf{Q}, \tag{10}$$

$$\mathbf{P}(g+1) = \mathbf{S}(g)(1 - 2\mathbf{M} \cdot \mathbf{Q}) + \mathbf{Q} \times \mathbf{P}(g+1) + \upsilon_h \nabla R(P(g))\mathbf{Q}, \tag{11}$$

$$\mathbf{P}(g+1) - \mathbf{Q} \times \mathbf{P}(g+1) = \mathbf{S}(g)(1 - 2\mathbf{M} \cdot \mathbf{Q}) + \upsilon_h \nabla R(P(g))\mathbf{Q}, \tag{12}$$

$$\mathbf{P}(g+1)(1 - \mathbf{Q}) = \mathbf{S}(g)(1 - 2\mathbf{M} \cdot \mathbf{Q}) + \upsilon_h \nabla R(P(g))\mathbf{Q}, \tag{13}$$

$$\mathbf{P}(g+1) = \frac{1}{(1 - \mathbf{Q})} [\mathbf{S}(g)(1 - 2\mathbf{M} \cdot \mathbf{Q}) + \upsilon_h \nabla R(P(g))\mathbf{Q}], \tag{14}$$

where $(g+1)$ denotes the next iteration, $\mathbf{M}$ signifies the random vector within the range [0.5, 1], $\mathbf{S}(g)$ indicates the position vector of target, $\mathbf{P}(g)$ symbolizes the position vector of sea lion, $\mathbf{L}$ signifies the total distance between the target prey and sea lion, $\mathbf{Q}$ specifies the constraints that linearly reduces the range from 2 to 0 for each iterations.

$$\upsilon_h = \frac{|(P(g)-P(g-1))^H[\nabla R(P(g))-\nabla R(P(g-1))]|}{\|[\nabla R(P(g))-\nabla R(P(g-1))]\|^2}. \tag{15}$$

(4) *Vocalization stage*

Sea lions are interacted with further sea lions by vocalizations while they hunt and chase as a subparts. When sea lion identifies a prey, the sea lion calls remaining members to encircle and assault the prey that is specified as

$$P_{\text{leader}} = |(\mathbf{C}_1 + (1+\mathbf{C}_2))/\mathbf{C}_2|, \tag{16}$$

where C_1 and C_2 indicate speed of sounds in water and air and speed of sea lion's sound represented as P_{leader}. Therefore, speed of sound in water is given as

$$\mathbf{C}_1 = \text{Sin}\,\theta. \tag{17}$$

In air, speed of sound is expressed as

$$\mathbf{C}_2 = \text{Sin}\,\phi. \tag{18}$$

(5) *Attacking stage*

To design hunting character of sea lions, two various segments, like dwindling encircling and circle upgrading are established and are explained as

(a) *Dwindling encircling strategy*

This dwindling encircling character depends on $\mathbf{Q}$ which supports sea lion leader to move along track of prey and encircle them.

(b) *Circle updating position*

Sea lions chase bait ball of fishes initializing from boundaries that can be expressed as

$$\mathbf{P}(g+1) = |\mathbf{S}(g) - \mathbf{P}(g)| \cdot \text{Cos}(2\pi n) + \mathbf{S}(g). \tag{19}$$

Consider $\mathbf{S}(g) > \mathbf{P}(g)$

$$\mathbf{P}(g+1) = (\mathbf{S}(g) - \mathbf{P}(g)) \cdot \text{Cos}(2\pi n) + \mathbf{S}(g), \tag{20}$$

$$\mathbf{P}(g+1) = \mathbf{S}(g) \cdot (\text{Cos}(2\pi n) + 1) - \mathbf{P}(g)\text{Cos}(2\pi n). \tag{21}$$

By substituting (9) in (21), the equation becomes

$$\mathbf{P}(g+1) = \mathbf{S}(g) \cdot (\text{Cos}(2\pi n) + 1) - (\mathbf{P}(g+1) + \upsilon_h \nabla R(P(g))) \cdot \text{Cos}(2\pi n), \tag{22}$$

$$\mathbf{P}(g+1) = \mathbf{S}(g) \cdot (\mathrm{Cos}(2\pi n) + 1) - \mathbf{P}(g+1) \cdot \mathrm{Cos}(2\pi n) - \upsilon_h \nabla R(P(g)) \cdot \mathrm{Cos}(2\pi n), \tag{23}$$

$$\mathbf{P}(g+1) + \mathbf{P}(g+1) \cdot \mathrm{Cos}(2\pi n) = \mathbf{S}(g) \cdot (\mathrm{Cos}(2\pi n) + 1) - \upsilon_h \nabla R(P(g)) \cdot \mathrm{Cos}(2\pi n), \tag{24}$$

$$\mathbf{P}(g+1)(1 + \mathrm{Cos}(2\pi n)) = \mathbf{S}(g) \cdot (\mathrm{Cos}(2\pi n) + 1) - \upsilon_h \nabla R(P(g)) \cdot \mathrm{Cos}(2\pi n), \tag{25}$$

$$\mathbf{P}(g+1) = \frac{1}{1 + \mathrm{Cos}(2\pi n)} [\mathbf{S}(g) \cdot (\mathrm{Cos}(2\pi n) + 1) - \upsilon_h \nabla R(P(g)) \cdot \mathrm{Cos}(2\pi n)], \tag{26}$$

where $\mathrm{Cos}(2\pi n)$ denotes behavior of sea lions swimming in circle shaped path around the prey, the distance between the finest solution and search agent is represented as $|\mathbf{S}(g) - \mathbf{P}(g)|$, $||$ denotes absolute value, and n represents the random integer that lies between $[-1, 1]$.

(6) *Prey searching*

Sea lions location is restructured by the optimal search agent in the exploitation step, whereas by consisting the arbitrarily selected sea lion, the sea lions position is updated in the exploration stage. However, the equation can be expressed as

$$\mathbf{L} = |2\mathbf{M} \cdot \mathbf{P}_{\mathrm{rand}}(g) - \mathbf{P}(g)|, \tag{27}$$

where $\mathbf{P}_{\mathrm{rand}}(g)$ represents randomly selected sea lion from the existing population.

$$\mathbf{P}(g+1) = \mathbf{P}_{\mathrm{rand}}(g) - \mathbf{L} \cdot \mathbf{Q}. \tag{28}$$

(7) *Feasibility evaluation*

Using fitness function, optimal result is computed such that the optimal rate of fitness is considered as optimum solution.

(8) *Termination*

The above phases are repetitively executed till the best solution is attained. The pseudo code of GDSLO is described in Table 1.

GDSLO + U-Net is efficient for extracting road surface and generated road segmented output is signified as S_i.

3.4. *Road edge and centerline detection using FCN*

Road edge detection is the process for extracting the boundaries of roads with single pixel-width. Thereafter, a centerline extraction method is accomplished for extracting the final network of road centerlines. Moreover, road edge detection and road centerline extraction are performed using FCN[28] by considering the output of the road segmentation S_i and pre-processed output P_i. Every layer of data in a ConvNet is a three-dimensional array of dimension $i \times j \times k$, where i and j represent spatial dimensions, and d denotes the dimension of the channel or attribute. Initial

Table 1. Pseudo code of established GDSLO algorithm.

1	Input: **Solution set** P, **maximum iterations** $g_{\max}$
2	**Output:** Best solution $\mathbf{P}(g+1)$
3	**Begin**
4	Solution initialization
5	Choose $\mathbf{P}_{\text{rand}}(g)$
6	Estimate fitness for each search agent using Eq. (3)
7	$\mathbf{P}^*$ denotes finest candidate search agent with best fitness
8	**If** $g < g_{\max}$
9	Estimate $\mathbf{P}_{\text{leader}}$ by Eq. (16)
10	**If** ($\mathbf{P}_{\text{leader}} < 0.25$)
11	**If** ($\lvert Q \rvert < 1$)
12	Upgrade prevailing location search agent based on Eq. (26)
13	**Else**
14	select random search agent $\mathbf{P}_{\text{rand}}(h)$
15	Upgrade location of current search agent based on Eq. (14)
16	**End if**
17	Upgrade location of prevailing search agent based on Eq. (28)
18	**End if**
19	**If** search agent does not belong to any $\mathbf{P}_{\text{leader}}$
20	Go to step 9
21	**Else**
22	Estimate the search agent fitness based on Eq. (3)
23	Update $\mathbf{P}$, if optimum solution exists
24	**Return P**
25	**End if**
26	**End if**
27	**End**

layer comprises of image with pixel dimension $i \times j$, and e color channels. The position of higher layer corresponds to the position of the image they are path-connected to, and are known as receptive fields. Moreover, ConvNets are constructed according to translation invariance. The components of ConvNets include convolution layer, pooling layer, and activation functions. These elements operate on regions of local input and are dependent on the relative spatial coordinates.

Consider a data vector x_{ij} at position (i, j) in a specific layer, and y_{ij} for the subsequent layer. However, the output computed using these functions is expressed as

$$y_{ij} = t_{sk}(\{x_{ki+\tau i, kj+\tau j}\} 0 \leq \tau i, \tau j \leq s), \tag{29}$$

where the kernel size is denoted as s, k signifies the sub-sampling factor or strides, t_{sk} evaluates the type of the layer: average pooling or convolutional matrix multiplication, activation function with element-wise nonlinearity or spatial max with maxpooling, and so on for other kinds of layers.

Under composition, the functional form is maintained with strides and kernel dimension by satisfying the transformation rule, which is given as

$$t_{sk} \circ b_{s'k'} = (t \circ b)s' + (s-1)k', kk'. \tag{30}$$

As the nonlinear general function is computed by the general deep net, a net with simply the layers of this form measures a nonlinear filter, which is called as the FCN or deep filter.

An FCN normally works on an input of any dimension, and generates a result of equivalent spatial sizes. In addition, an FCN with real-valued loss function specifies a task. If the loss function is the overall quantity over the spatial sizes of the last layer, $r(x;\vartheta) = \sum_{ij} r'(x_{ij};\vartheta)$ its gradient is the total sum computed over the gradients of every spatial elements. Hence, the SGD on measured r on overall images is same as the SGD on r', considering all the receptive fields of final layer as a minibatch.

Furthermore, when the receptive fields considerably overlap, the backpropagation and feedforward calculation are considered to be more effectual when computed layer-by-layer over a full image as a replacement for independent patch-by-patch. Finally, the detected road edge-centerline output is represented as E_i.

4. Results and Discussion

This segment establishes the execution consequences of established GDSLO-based U-Net considering the calculation procedures, specifically precision, recall, and F-measure.

4.1. *Experimental setup*

The implementation of the proposed GDSLO-enabled U-Net is performed in PYTHON tool with PyTorch using RoadNet dataset (Dataset-1)[25] and Massachusetts road dataset (Dataset-2)[26] with PC having 2GB RAM windows 10 OS, and Intel i3 core processor. Table 2 shows the experimental parameters of the devised GDSLO-enabled U-Net.

4.2. *Dataset description*

The two different datasets used for the experimentation are explained as follows:

Dataset-1: This dataset[25] collects various typical urban regions of Canada, Ottawa, from Google Earth. In this dataset, the images include 21 regions about 8 km with 0.21-m three-dimensional resolution per pixel. Here, the road exteriors, road borders, and road midways for every image is manually annotated. The width of the road ranges from 10 to 80 pixels and this dataset learns multi-scale and multi-level features to deal occlusions with and shadows issues.

Table 2. Experimental parameters.

Parameter	Value
Learning rate	0.0001
Momentum	0.9
Step size	5
Gamma	0.1
Number of epochs	200

Dataset-2: This dataset[26] comprises images of $1500 * 1500$ pixels covering a region of $500\,\text{km}^2$ at a $1.2\,\text{m/pixel}$ resolution. In this dataset, the labels and the imageries are publicly available. However, the labels are scraped from OpenStreetMap and not collected or validated separately. The dataset was augmented by applying vertical, horizontal flips and also by varying the intensities at different degrees during the time of training.

4.3. *Performance metrics*

The performance of established GDSLO-enabled U-Net technique is examined in terms of three evaluation standards, such as precision, recall, and F1-measure, respectively.

Precision: Precision is a measure which defines the fraction of accurately computed outcomes, and is expressed as

$$P = \frac{a}{a+b}, \tag{31}$$

where a denotes true positives, and b specifies false positives.

***Recall*:** It defines as ratio of the true positives to addition of true positives and false negatives, and can be expressed as

$$R = \frac{a}{a+c}, \tag{32}$$

where a represents the true positives, and c signifies the false negatives.

***F1-measure*:** F1-measure estimates mean difference between the precision and recall, and can be demonstrated as

$$\text{FM} = 2 * \left(\frac{P * R}{P + R}\right), \tag{33}$$

where P represents the precision, and R signifies the recall.

4.4. *Experimental outcomes*

The implementation results of established GDSLO-based U-Net for road detection is depicted below as follows. Figure 3(a) shows the original centerline and edge extracted image, Fig. 3(b) depicts the detected image of centerline and road extraction, Fig. 3(c) signifies the ground truth images.

4.5. *Comparative methods*

The various comparative techniques taken for the assessment are CNN,[1] Cascaded CNN,[3] ICNet,[4] LWNet,[5] FP network, ScRoadExtractor,[8] SGD-based U-Net, SLnO-based U-Net, and developed GDSLO-based U-Net, respectively.

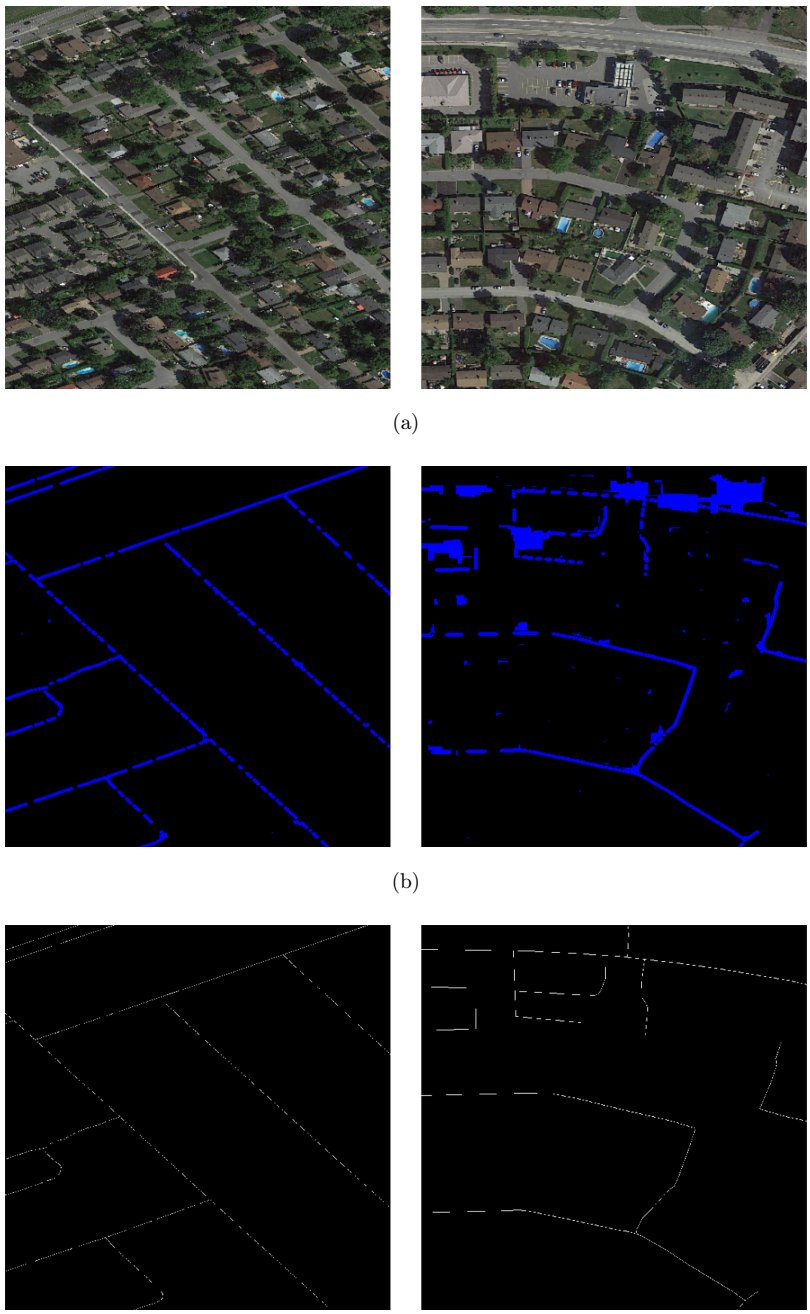


Fig. 3. Experimental outcomes of the developed method (a) original image of centerline and edge extraction image, (b) detected image of centerline and edge extraction (c) ground truth images.

4.6. *Comparative assessment*

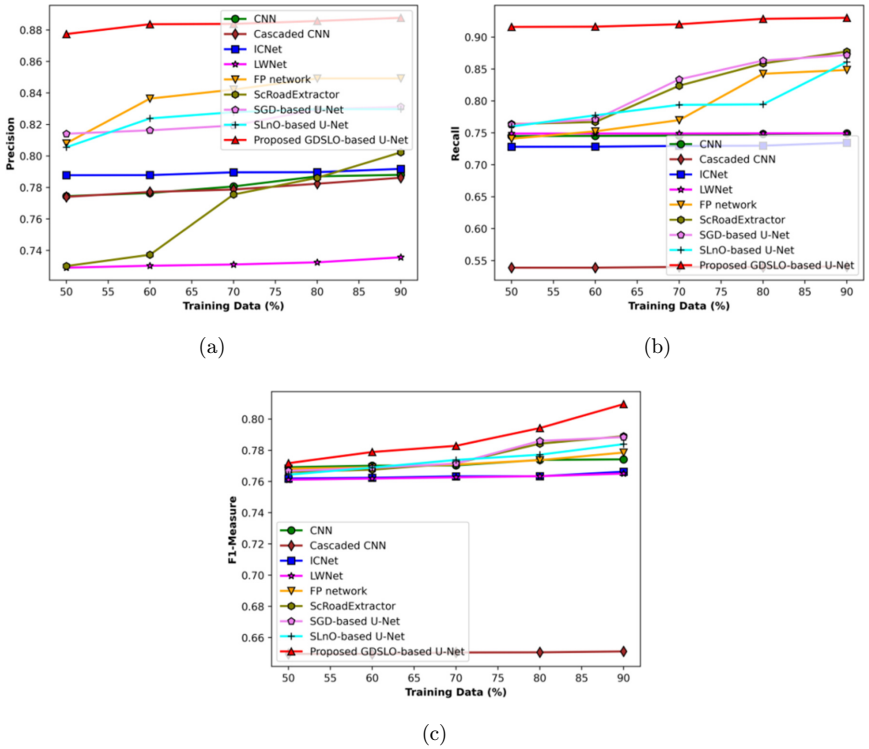
This section explains the comparative assessment of proposed GDSLO-enabled U-Net method based on dataset-1 and dataset-2 by varying the training data in terms of evaluation metrics.

4.6.1. *Analysis using dataset-1*

This segment illustrates the testing performed using dataset-1 for road surface extraction, road edge detection, and road centerline detection by considering the evaluation metrics.

(a) Analysis based on road surface extraction

Figure 4 displays the analysis of the established technique by means of dataset-1 by altering the training data relating to the estimation standards. Figure 4(a) describes analysis depending upon the precision. Analyzing the training data 50%, the precision attained by the established GDSLO-based U-Net is 0.877, however, precision value measured by the current techniques, such as CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, is 0.774, 0.774, 0.787, 0.729, 0.808, 0.730, 0.814, and 0.806. The assessment by means



(a) (b)

(c)

Fig. 4. Analysis with regard to road surface extraction (a) Precision, (b) Recall, and (c) F1-measure.

of recall measure is shown in Fig. 4(b). For training data 60%, the established GDSLO+ U-Net measured a recall value of 0.916, while prevailing methods, namely, CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, computed the recall value of 0.745, 0.538, 0.728, 0.748, 0.752, 0.767, 0.771, and 0.777. The analysis is done by using F1-measure depicted in Fig. 4(c). The F1-measure value attained by the CNN is 0.770, Cascaded CNN is 0.650, ICNet is 0.763, LWNet is 0.762, FP network is 0.771, ScRoadExtractor is 0.772, SGD-based U-Net is 0.772, SLnO-based U-Net is 0.774, and developed GDSLO-based U-Net is 0.782 for training data 70%.

(b) Analysis based on road edge detection

The analysis of the developed method is done by using dataset-1 as shown in Fig. 5. The analysis by means of precision can be illustrated in Fig. 5(a). When training data becomes 60%, the developed GDSLO-based U-Net measured a precision of 0.800. On the other hand, the precision value is measured using the existing techniques, such as CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net is 0.799, 0.799, 0.799, 0.799, 0.800, 0.800, 0.800, and 0.800. Figure 5(b) shows the calculation depending on recall measure with 70% of

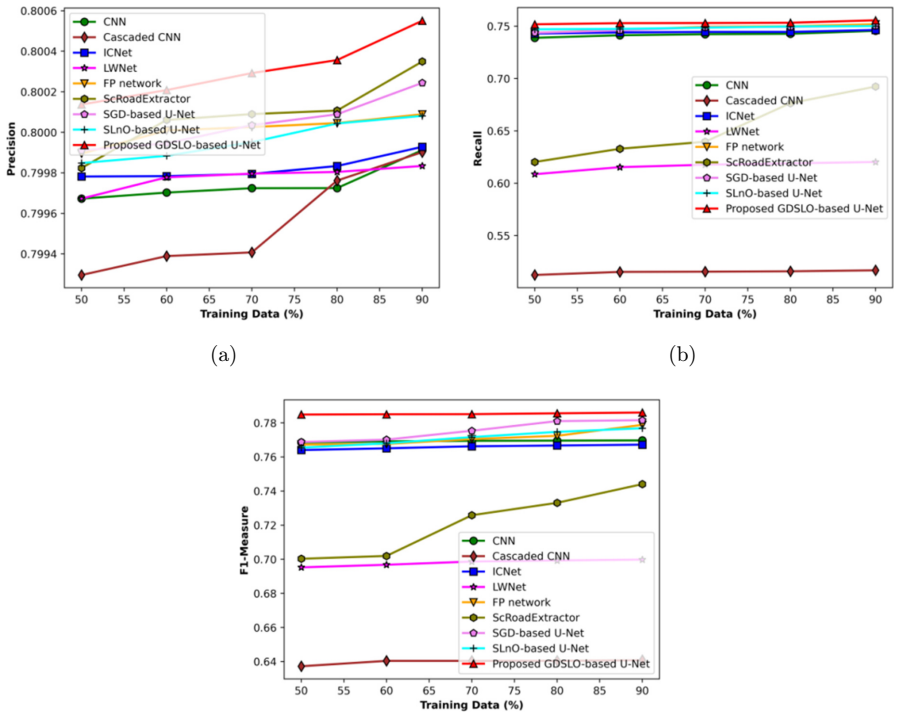
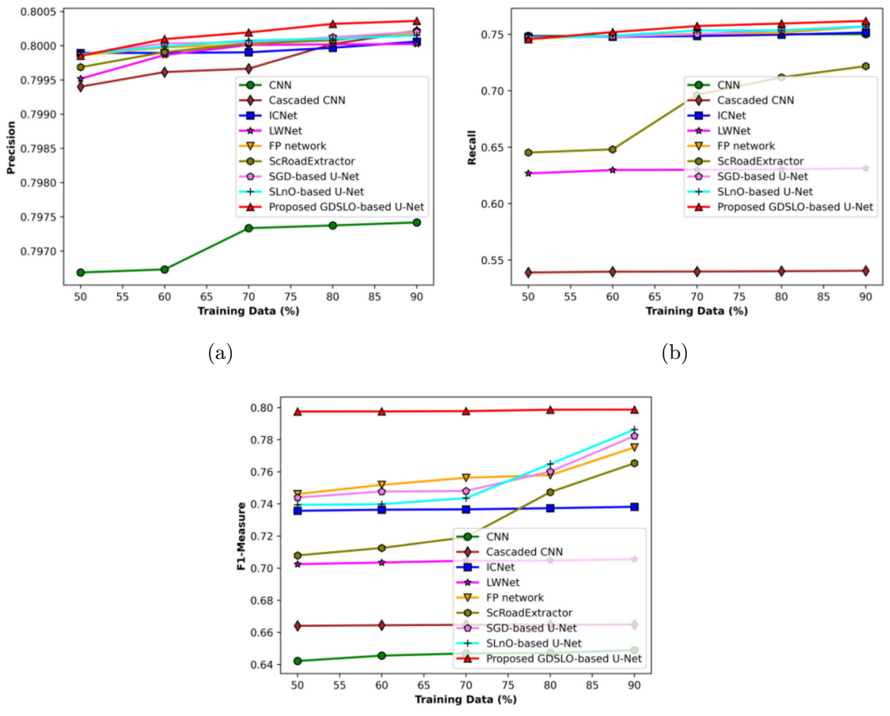


Fig. 5. Analysis with respect to road edge detection (a) Precision, (b) Recall, and (c) F1-measure.

training data, recall attained by CNN is 0.741, Cascaded CNN is 0.515, ICNet is 0.744, LWNet is 0.615, FP network is 0.749, ScRoadExtractor is 0.640, SGD-based U-Net is 0.749, SLnO-based U-Net is 0.749, and developed GDSLO-based U-Net is 0.752. The analysis can be made by using F1-measure as revealed in Fig. 5(c). With 80% of training data, F1-measure is obtained by the established GDSLO-based U-Net which is 0.785. Meanwhile, the F1-measure obtained by prevailing methods, including CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, is 0.769, 0.640, 0.766, 0.698, 0.772, 0.733, 0.781, and 0.775, respectively.

(c) Analysis based on road centerline detection

The analysis of the established technique depending upon the dataset-1 by shifting the training data is revealed in Fig. 6. The evaluation based on precision is shown in Fig. 6(a). With 50% of training data, GDSLO+U-Net evaluates a precision of 0.7998, while prevailing methods, namely, CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, computed a precision value of 0.7966, 0.7994, 0.7998, 0.7995, 0.7998, 0.7996, 0.7998, and 0.7998, respectively. Figure 6(b) demonstrates the analysis based on recall measure. To consider the training data 60%, value of recall found by the developed



(a) (b)

(c)

Fig. 6. Analysis with respect to road centerline detection (a) Precision, (b) Recall, and (c) F1-measure.

GDSLO+U-Net is 0.751, whereas the recall value measured by the current techniques, such as CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, is 0.748, 0.539, 0.747, 0.629, 0.748, 0.648, 0.747, and 0.748, respectively. The analysis is done with regard to F1-measure as depicted in Fig. 6(c). For training data of 70%, F1-measure is calculated by CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, SLnO-based U-Net, and developed GDSLO-based U-Net is 0.646, 0.664, 0.736, 0.704, 0.756, 0.719, 0.748, 0.743, and 0.797, respectively.

(d) Analysis based on testing time and training time using dataset-1

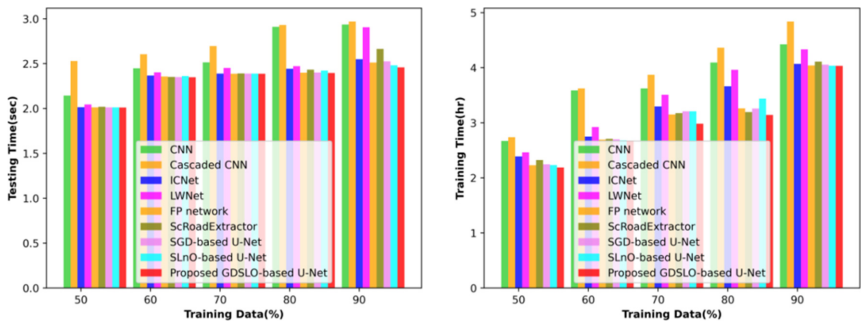
Figure 7 establishes the analysis based on testing time and training time. Figure 7(a) represents the analysis depending on testing time. For 60% of training data, the value of testing time calculated by CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, and developed GDSLO+U-Net is 2.447, 2.606, 2.366, 2.402, 2.352, 2.351, 2.362, and 2.349 s, respectively. The assessment depending on training time is depicted in Fig. 7(b). In varying training data of 70%, the established GDSLO-based U-Net measured a training time of 2.984 s, whereas the training time computed by the CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, is 3.625, 3.872, 3.297, 3.509, 3.152, 3.175, 3.210, and 3.210 s, respectively.

4.6.2. *Analysis using dataset-2*

This phase elucidates the analysis performed using dataset-2 depending on the evaluation measures by shifting training data percentage.

(a) Analysis based on road detection

Figure 8 explains the analysis of the advanced approach by means of dataset-2. The analysis is made with precision which is illustrated in Fig. 8(a). When the training data is 50%, the developed GDSLO-based U-Net measured a precision of 0.674. On



(a) (b)

Fig. 7. Analysis of the developed technique using dataset-1 considering (a) Testing time (b) Training time.

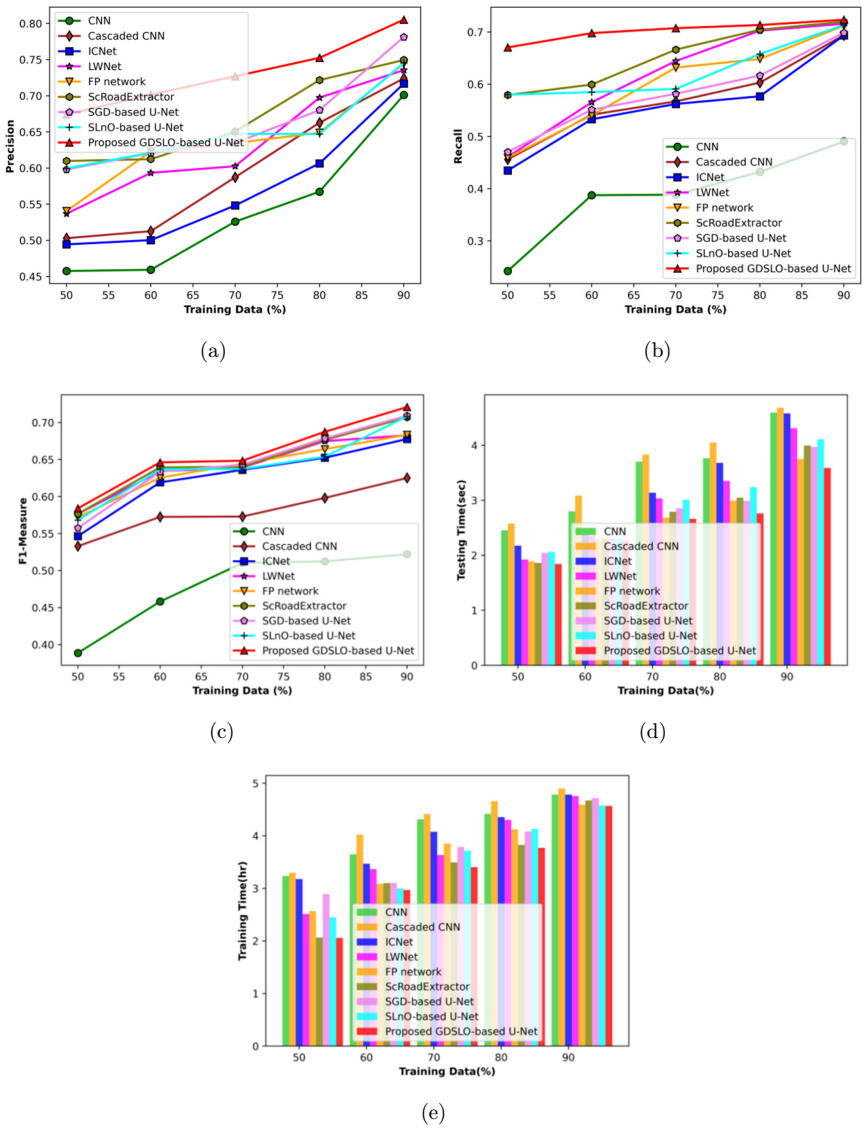


Fig. 8. Analysis with respect to road detection (a) Precision, (b) Recall, (c) F1-measure, (d) Testing time, and (e) Training time.

the other hand, the value determined by the prevailing methods, such as CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, is 0.457, 0.502, 0.494, 0.536, 0.541, 0.610, 0.598, and 0.599, respectively. Figure 8(b) presents the assessment based on recall measure. For 60% of training data, recall deducted by CNN is 0.387, Cascaded CNN is 0.541, ICNet is 0.532, LWNet is 0.565, FP network is 0.540, ScRoadExtractor is 0.599, SGD-based U-Net is 0.551, SLnO-based U-Net is 0.584, and developed GDSLO-based U-Net is

0.697. The analysis is done by using F1-measure which is revealed in Fig. 8(c). With sample data of 70%, the value of F1-measure obtained by the established GDSLO+U-Net is 0.648. Meanwhile, the F1-measure attained by the present methods, such as CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net is 0.510, 0.572, 0.636, 0.639, 0.644, 0.640, 0.643, and 0.637, respectively. Figure 8(d) illustrates the analysis based on testing time and training time. Figure 8(a) represents the analysis based on testing time. For sample data of 60%, the testing time value computed by CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, and developed GDSLO-based U-Net is 2.799, 3.086, 2.381, 2.356, 2.290, 2.304, 2.301, 2.305, and 2.258 s, respectively. The evaluation based on training time is displayed in Fig. 8(e). By varying training data of 70%, advanced GDSLO-based U-Net measured a training time of 3.402 s, whereas the training time computed by the CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, 4.313, 4.412, 4.076, 3.632, 3.853, 3.493, 3.789 and 3.717 s, respectively.

4.7. *Comparative discussion*

Table 3 describes the relative evaluation of the proposed GDSLO-enabled U-Net method in contrast with various existing systems, such as CNN, Cascaded CNN, ICNet, LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, using dataset-1 and dataset-2 by varying the training data of 90%.

Table 3. Relative discussion of the established technique using database-1 and database-2.

Methods/Metrics		Precision	Recall	F1-measure
			Dataset-1	
Road surface extraction	CNN	0.788	0.749	0.774
	Cascaded CNN	0.786	0.541	0.651
	ICNet	0.792	0.735	0.766
	LWNet	0.736	0.749	0.765
	FP network	0.849	0.849	0.779
	ScRoadExtractor	0.802	0.877	0.789
	SGD-based U-Net	0.831	0.872	0.788
	SLnO-based U-Net	0.830	0.861	0.784
	Proposed GDSLO-based U-Net	**0.888**	**0.930**	**0.810**
Road edge detection	CNN	0.799	0.746	0.770
	Cascaded CNN	0.799	0.517	0.641
	ICNet	0.799	0.746	0.767
	LWNet	0.799	0.620	0.700
	FP network	0.800	0.752	0.779
	ScRoadExtractor	0.800	0.692	0.744
	SGD-based U-Net	0.800	0.750	0.782
	SLnO-based U-Net	0.800	0.750	0.777
	Proposed GDSLO-based U-Net	**0.801**	**0.756**	**0.786**

Table 3. (*Continued*)

Methods/Metrics		Precision	Recall	F1-measure
Road centerline detection	CNN	0.797	0.750	0.649
	Cascaded CNN	0.800	0.541	0.665
	ICNet	0.800	0.752	0.738
	LWNet	0.800	0.631	0.706
	FP network	0.800	0.757	0.775
	ScRoadExtractor	0.800	0.722	0.765
	SGD-based U-Net	0.800	0.757	0.782
	SLnO-based U-Net	0.800	0.757	0.786
	Proposed GDSLO-based U-Net	**0.800**	**0.762**	**0.799**
Road detection			***Dataset-2***	
	CNN	0.701	0.490	0.522
	Cascaded CNN	0.724	0.694	0.625
	ICNet	0.717	0.693	0.678
	LWNet	0.735	0.716	0.683
	FP network	0.744	0.712	0.683
	ScRoadExtractor	0.749	0.720	0.707
	SGD-based U-Net	0.781	0.699	0.709
	SLnO-based U-Net	0.746	0.713	0.708
	Proposed GDSLO-based U-Net	**0.805**	**0.723**	**0.721**

For the road surface extraction, the developed GDSLO-based U-Net measured a precision, recall, and F1-measure of 0.887, 0.930, and 0.809 using dataset-1. Likewise, the developed GDSLO-based U-Net calculated a precision value of 0.800, recall value 0.755 and F1-measure value of 0.786 for road edge detection using dataset-1. In addition, the precision, recall and F1-measure value achieved by the developed GDSLO-based U-Net for road centerline detection is 0.800, 0.761, and 0.798 using dataset-1. Furthermore, the established GDSLO+U-Net accomplished a precision value of 0.805, recall value of 0.723 and F1-measure value of 0.720 for road detection using dataset-2.

Table 4 elucidates the relative discussion of the established method in contrast with the prevailing methods, such as CNN, Cascaded CNN, ICNet, and LWNet, FP network, ScRoadExtractor, SGD-based U-Net, and SLnO-based U-Net, using dataset-1 and dataset-2 for testing time and training time. The developed

Table 4. Comparative discussion of the established technique based on dataset-1 and dataset-2 for testing time and training time.

Metrics	CNN	Cascaded CNN	ICNet	LWNet	FP network	ScRoad extractor	SGD-based U-Net	SLnO-based U-Net	Proposed GDSLO-based U-Net
Testing time (s)	2.936	2.969	2.549	2.905	2.513	2.664	2.525	2.482	**2.458**
Training time (s)	4.784	4.900	4.783	4.754	4.041	4.112	4.055	4.036	**4.566**
Testing time (s)	4.594	4.684	4.580	4.311	3.751	3.998	3.970	4.111	**3.586**
Training time (s)	4.784	4.900	4.783	4.754	4.593	4.673	4.717	4.574	**4.566**

GDSLO-based U-Net achieved a testing time of 2.458 s using dataset-1 and 3.586 s using dataset-2. Likewise, the training time computed by the developed GDSLO-based U-Net is 3.586 s and 4.566 using dataset-2.

5. Conclusion

Automatically extracting the road surfaces from distant sensed images performs a vital role in geospatial data integration, vehicle navigation, urban design, and intelligent transportation system. On the other hand, it is dull and time consuming to distinct the roads from HR remote sensing images manually. This research offers a novel deep learning-enabled road segmentation and edge-midway extraction model to extract the road surfaces using GDSLO-based U-Net. The input image is pre-processed and road segmentation is done using U-Net for extracting the road surfaces effectively. However, the training data of U-Net is completed by means of established optimization algorithm, named as GDSLO. The established GDSLO procedure is intended by combination of SGD and SLnO. After extracting road surfaces, road edge detection is done using FCN for extracting single-pixel width boundaries of roads and road centerline detection is performed using FCN for representing the road networks. Furthermore, the developed GDSLO-based U-Net outperformed numerous existing methods and achieved effective performance by considering precision, recall, and F1-measure with maximum values of 0.887, 0.930, and 0.809, respectively. However, the determination of wrong road topology and unclosed boundary are the key issues in the proposed technique. This will be considered in our future work by exploring the loss function and measure the geometric similarity and topology.

References

1. Y. Liu, J. Yao, X. Lu, M. Xia, X. Wang and Y. Liu, "RoadNet: Learning to comprehensively analyze road networks in complex urban scenes from high-resolution remotely sensed images," *IEEE Trans. Geosci. Remote Sens.* **57**(4), 2043–2056 (2018).
2. R. Liu, Q. Miao, J. Song, Y. Quan, Y. Li, P. Xu and J. Dai, "Multiscale road centerlines extraction from high-resolution aerial imagery," *Neurocomputing* **329**, 384–396 (2019).
3. 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 Trans. Geosci. Remote Sens.* **55**(6), 3322–3337 (2017).
4. R. Senchuri, A. Kuras and I. Burud, "Machine learning methods for road edge detection on fused airborne hyperspectral and LIDAR Data," in *11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)* (Amsterdam, Netherlands, 2021), pp. 1–5.
5. Z. Chen and Z. Chen, "RBnet: A deep neural network for unified road and road boundary detection," in *Int. Conf. Neural Information Processing* (Springer, Cham, 2017), pp. 677–687.
6. Y. Wei, K. Zhang and S. Ji, "Simultaneous road surface and centerline extraction from large-scale remote sensing images using CNN-based segmentation and tracing," *IEEE Trans. Geosci. Remote Sens.* **58**(12), 8919–8931 (2020).

7. P. Shamsolmoali, M. Zareapoor, H. Zhou, R. Wang and J. Yang, "Road segmentation for remote sensing images using adversarial spatial pyramid networks," *IEEE Trans. Geosci. Remote Sens.* **59**(6), 4673–4688 (2020).
8. Y. Wei and S. Ji, "Scribble-based weakly supervised deep learning for road surface extraction from remote sensing images," *IEEE Trans. Geosci. Remote Sens.* **99**, 1–12 (2021).
9. R. Lian, W. Wang, N. Mustafa and L. Huang, "Road extraction methods in high-resolution remote sensing images: A comprehensive review," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **13**, 5489–5507 (2020).
10. S. Wang, H. Yang, Q. Wu, Z. Zheng, Y. Wu and J. Li, "An improved method for road extraction from high-resolution remote-sensing images that enhances boundary information," *Sensors* **20**(7), 2064 (2020).
11. L. Li, D. Zhang, S. Ying and Y. Li, "Recognition and reconstruction of zebra crossings on roads from mobile laser scanning data," *ISPRS Int. J. Geo-Inf.* **5**(7), 125 (2016).
12. H. Guan, J. Li, Y. Yu, Z. Ji and C. Wang, "Using mobile LiDAR data for rapidly updating road markings," *IEEE Trans. Intell. Transport. Syst.* **16**(5), 2457–2466 (2015).
13. S. Kaur and S. Baghla, "Automatic road detection of satellite images-A survey," *Int. J. Comput. Appl. Inf. Technol.* **3**(2), 32–34 (2013).
14. X. Li, S. Zhang, X. Pan, P. Dale and R. Cropp, "Straight road edge detection from high-resolution remote sensing images based on the ridgelet transform with the revised parallel-beam Radon transform," *Int. J. Remote Sens.* **31**(19), 5041–5059 (2010).
15. L. Xu, T. Jun, Y. Xiang, C. JianJie and G. LiQian, "The rapid method for road extraction from high-resolution satellite images based on USM algorithm," in *Int. Conf. Image Analysis and Signal Processing*, IEEE, Manhattan, New York, pp. 1–6.
16. L. Wang, Q. Qin, S. Du, D. Chen and J. Tao, "Road extraction from remote sensing image based on multi-resolution analysis," in *Int. Symp. Remote Sensing of Environment* (Tshwane, South Africa, 2005).
17. R. Liu, J. Song, Q. Miao, P. Xu and Q. Xue, "Road centerlines extraction from high resolution images based on an improved directional segmentation and road probability," *Neurocomputing* **212**, 88–95 (2016).
18. X. Hu, Z. Zhang and C. V. Tao, "A robust method for semi-automatic extraction of road centerlines using a piecewise parabolic model and least square template matching," *Photogramm. Eng. Remote Sens.* **70**(12), 1393–1398 (2004).
19. C. Tao, J. Qi, Y. Li, H. Wang and H. Li, "Spatial information inference net: Road extraction using road-specific contextual information," *ISPRS J. Photogramm. Remote Sens.* **158**, 155–166 (2019).
20. P. K. Soni, N. Rajpal and R. Mehta, "Semiautomatic road extraction framework based on shape features and LS-SVM from high-resolution images," *J. Ind. Soc. Remote Sens.* **48**(3), 513–524 (2020).
21. X. Gao, X. Sun, Y. Zhang, M. Yan, G. Xu, H. Sun, J. Jiao and K. Fu, "An end-to-end neural network for road extraction from remote sensing imagery by multiple feature pyramid network," *IEEE Access* **6**, 39401–39414 (2018).
22. A. Wulamu, Z. Shi, D. Zhang and Z. He, "Multiscale road extraction in remote sensing images," *Comput. Intell. Neurosci.* (2019).
23. R. Masadeh, B. A. Mahafzah and A. Sharieh, Sea lion optimization algorithm, *Int. J. Adv. Comput. Sci. Appl.* **10**(5), 388–395 (2019).
24. S. V. Kumar and C. Nagaraju, "T2FCS filter: Type 2 fuzzy and cuckoo search-based filter design for image restoration," *J. Vis. Commun. Image Represent.* **58**, 619–641 (2019).
25. RoadNet dataset (2021), https://github.com/yhlleo/RoadNet.

26. Massachusetts dataset (2021), https://www.cs.toronto.edu/∼vmnih/data/mass_roads/train/sat/index.html.
27. O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Int. Conf. Medical Image Computing and Computer-Assisted Intervention* (Springer International Publishing, Switzerland, 2015), pp. 234–241.
28. J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (2015), pp. 3431–3440.
29. V. Rupapara, F. Rustam, H. F. Shahzad, A. Mehmood, I. Ashraf and G. S. Choi, "Impact of SMOTE on imbalanced text features for toxic comments classification using RVVC model," *IEEE Access* **9**, 78621–78634 (2021).
30. D. K. Shende and S. S. Sonavane, "CrowWhale-ETR: CrowWhale optimization algorithm for energy and trust aware multicast routing in WSN for IoT applications," *Wirel. Netw.* **26**(6), 4011–4029 (2020).
31. D. K. Shende, Y. Angal and S. S. Sonavane, "A comprehensive survey of the routing schemes for IoT applications," *Scal. Comput. Pract. Exp.* **21**(2), 203–216 (2020).
32. G. R. Bojja and L. S. Ambati, "A novel framework for crop pests and disease identification using social media and AI," in *Proc. MWAIS* (2020), pp. 1–4.
33. P. Jegatheeswari and T. A. Deepa, "Fuzzy weighted least square filter for pansharpening in satellite images," *Multimed. Res.* **2**(1), 17–22 (2019).
34. S. Vinusha, "Secret image sharing and steganography using Haar wavelet transform," *Multimed. Res.* **2**(2), 28–34 (2019).
35. E. Regentova, D. Yao, S. Latifi and J. Zheng, "Image segmentation using Ncut in the wavelet domain," *Int. J. Image Graph.* **6**(4), 569–582 (2006).
36. J. Lin, B. Peng and T. Li, "A learning-based framework for supervised and unsupervised image segmentation evaluation," *Int. J. Image Graph.* **14**(3), 1450014 (2014).
37. J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang and H. Lu, "Dual attention network for scene segmentation," in *IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 3146–3154.
38. Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei and W. Liu, "CCNet: criss-cross attention for semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Computer Vision (ICCV)* (2019), pp. 603–612.
39. X. Hu, K. Yang, L. Fei and K. Wang, "ACNET: Attention based network to exploit complementary features for RGBD semantic segmentation," in *Proc. IEEE Int. Conf. Image Processing (ICIP)* (IEEE, Taipei, Taiwan, 2019).
40. S. Zheng, J. Lu, H. Zhao, X. Zhu, Z. Luo, Y. Wang, Y. Fu, J. Feng, T. Xiang, Philip, H. S. Torr and L. Zhang, "Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2021).
41. E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez and P. Luo, "SegFormer: Simple and efficient design for semantic segmentation with transformers," in *Proc. Advances in Neural Information Processing Systems 34 (NeurIPS 2021)* (2021).
42. J. Zhang, K. Yang, A. Constantinescu, K. Peng, K. Müller and R. Stiefelhagen, "Trans4Trans: efficient transformer for transparent object and semantic scene segmentation in real-world navigation assistance," *IEEE Trans. Intell. Transport. Syst.* 1–14 (2022).
43. L. Sun, J. Wang, K. Yang, K. Wu, X. Zhou, K. Wang and J. Bai, "Aerial-PASS: Panoramic annular scene segmentation in drone videos," *Eur. Conf. Mobile Robots (ECMR), in Proc. Eur. Conf. Mobile Robots (ECMR)* (IEEE, Bonn, Germany, 2021).

44. Q. Qian, R. Jin, J. Yi, L. Zhang and S. Zhu, "Efficient distance metric learning by adaptive sampling and mini-batch stochastic gradient descent (SGD)," *Mach. Learn.* **99**, 353–372 (2015).

Miral J Patel received her B.E. degree in Electronics and Communication from C. U. Shah Engineering College, Saurashtra University, Rajkot, Gujarat, India in 2002 and M.E. degree in Electronics and Communication System Engineering from the Dharmsinh Desai University, Nadiad, 2007. She is pursuing her Ph.D from the Atmiya University, Rajkot, Gujarat. She has an experience of more than 15 years in teaching. Presently, she is working as Assistant Professor in Electronics & Communication Department, Government Engineering College, Rajkot. Her areas of interest are deep learning,image processing and computer networking.

Dr. Ashish M. Kothari received his Bachelor and Master degrees from the Saurashtra University, Rajkot, Gujarat, India, in 2005, and 2010, respectively. He earned his Ph.D. degree from the JJT University Rajasthan India. He is currently working as an Professor at the Atmiya University, Rajkot India. His research focuses on Image Classification and deep neural networks. He has published many research papers. His research interests include deep learning architectures and their applications for Image Classification.

2350059

https://doi.org/10.1142/S0219467823500596

Preview Abstract Abstract | PDF/EPUB

No Access

Deep Learning-Enabled Road Segmentation and Edge-Centerline Extraction from High-Resolution Remote Sensing Images

Miral Jerambhai Patel and Ashish M. Kothari

2350058

https://doi.org/10.1142/S0219467823500584

Preview Abstract Abstract | PDF/EPUB

No Access

Severity Detection of Diabetic Retinopathy — A Review

M. Gargi and Anupama Namburu

International Journal of Image and Graphics | Online Ready

Deep Learning-Enabled Road Segmentation and Edge-Centerline Extraction from High-Resolution Remote Sensing Images

Miral Jerambhai Patel and Ashish M. Kothari

https://doi.org/10.1142/S0219467823500584 | **Cited by:** 0

< Previous Next >

Abstract

Nowadays, precise and up-to-date maps of road are of great significance in an extensive series of applications. However, it automatically extracts the road surfaces from high-resolution remote sensed images which will remain as a demanding issue owing to the occlusion of buildings, trees, and intricate backgrounds. In order to address these issues, a robust Gradient Descent Sea Lion Optimization-based U-Net (GDSLO-based U-Net) is developed in this research work for road outward extraction from High Resolution (HR) sensing images. The developed GDSLO algorithm is newly devised by the incorporation of Stochastic Gradient Descent (SGD) and Sea Lion Optimization Algorithm (SLnO) algorithm. Input image is pre-processed and U-Net is employed in road segmentation phase for extracting the road surfaces. Meanwhile, training data of U-

Submissions with an Editorial Office Decision for Author

Page: 1 of 1 (1 total completed submissions)

Results per page 25

Action	Manuscript Number	Title	Initial Date Submitted	Current Status	Date Final Disposition Set	Final Disposition
View Submission Author Response View Decision Letter View Attachments Send E-mail	IJIG-D-22-00024	Deep learning enabled road segmentation and edge-centerline extraction from high-resolution remote sensing images	Feb 01, 2022	Accept		

Page: 1 of 1 (1 total completed submissions)

Results per page 25

Source details

Feedback > Compare sources >

International Journal of Image and Graphics

Scopus coverage years: from 2001 to Present

Publisher: World Scientific

ISSN: 0219-4678 E-ISSN: 1793-6756

Subject area: Computer Science: Computer Graphics and Computer-Aided Design | Computer Science: Computer Science Applications | Computer Science: Computer Vision and Pattern Recognition

Source type: Journal

View all documents > | Set document alert | Save to source list | Source Homepage

CiteScore 2021
1.5

SJR 2021
0.236

SNIP 2021
0.639

CiteScore | CiteScore rank & trend | Scopus content coverage

Improved CiteScore methodology

CiteScore 2021 counts the citations received in 2018-2021 to articles, reviews, conference papers, book chapters and data papers published in 2018-2021, and divides this by the number of publications published in 2018-2021. Learn more >