

# Bioengineering and Geomatics: Automatic Brain Image Segmentation using Two-Stage Pipeline with SNN and Watershed Algorithm

VINCENZO BARRILE<sup>1\*</sup>, EMANUELA GENOVESE<sup>1</sup>, ELENA BARRILE<sup>2</sup>

<sup>1</sup>Department of Civil, Energy, Environmental and Materials Engineering (DICEAM),  
Mediterranea University of Reggio Calabria,  
Via Graziella Feo di Vito- 89124, Reggio Calabria,  
ITALY

<sup>2</sup>Vita-Salute San Raffaele University,  
Via Olgettina, 58, 20132, Milan,  
ITALY

*\*Corresponding Author*

**Abstract:** - Digital image processing holds an increasingly essential role in the medical domain. This study emphasizes the significance of researching and implementing methods aimed at the segmentation of critical image regions and potential noise reduction, which is indispensable for medical professionals in disease diagnosis. Consequently, the investigation of software solutions in this context can substantially enhance diagnostic accuracy. In particular, neurology stands as a medical field wherein imaging plays a substantial contributory role. In pursuit of an automated brain image segmentation approach, this paper centers its attention on a two-step pipeline methodology to address the segmentation challenges inherent in medical imaging. The proposed method incorporates the use of a Self-Normalizing Neural Network (SNN) for denoising and employs the Watershed algorithm, typically employed in Geomatics imagery, for segmentation. Encouraging results are obtained, with a segmentation performance, as measured by IoU, reaching a noteworthy value of 0.93 when compared with alternative segmentation software.

**Key-Words:** - Neurology, biomedicine, neural networks, watershed technique, segmentation

Received: June 25, 2022. Revised: September 15, 2023. Accepted: October 5, 2023. Published: October 12, 2023.

## 1 Introduction

Segmentation in medical and geomatic images can pose challenges due to image variations in contrast, noise, and brightness, making it difficult to distinguish regions accurately. Additionally, images may contain artifacts like blurs, shadows, irregular shapes, diversity between subjects, and other complexities, which prolong and complicate the segmentation process, [1], [2], [3]. As manual brain segmentation is time-consuming, and segmenting brain boundaries is complicated due to shadows and noise, [4], [5], [6], machine learning and artificial intelligence offer new segmentation techniques that can recognize patterns in images and enhance image quality while minimizing artifacts. There are several segmentation techniques commonly used in medical imaging:

1. **Thresholding Technique:** This method involves comparing each pixel's intensity value with a chosen threshold. If the pixel's intensity is lower than the threshold, it's assigned to the background;

otherwise, it's assigned to the target region. The result is a binary image, where pixels below the threshold are set to 0, and those above it are set to 1. However, this technique can only generate two classes and does not consider spatial characteristics.

2. **Region-Based Segmentation:** Region-based algorithms partition the image into similar regions based on predefined criteria. One approach is "region growing," where initial regions are defined, and pixels are added to them if their intensity is similar to the region's average value. Another method is "region splitting and merging," which assumes the image is initially a single region and divides it into smaller regions if needed. Adjacent regions that meet certain criteria are then merged.

3. **Edge Detection:** This method focuses on recognizing contours in an image. It results in a binary image where contours are assigned a value of 1, while the background is set to 0. Edge detection relies on identifying changes in image intensity and often uses derivative filters to estimate pixel

gradients. However, merely detecting edges may not be sufficient to recognize significant regions in an image, as many edges may be incomplete or intersect.

4. Neural Networks: Neural network-based segmentation is a departure from conventional algorithms. It represents an image as a weighted graph, where nodes correspond to one or more pixels, and edge weights indicate similarity between adjacent pixels. Various algorithms can effectively partition these nodes to achieve segmentation, [7], [8], [9].

Each of these segmentation techniques has its strengths and limitations, and the choice of method depends on the specific characteristics of the images and the desired results in medical imaging applications. In this context, this paper introduces a new technique in this field: the watershed algorithm, commonly used in geomatics for orographic structure segmentation. However, the watershed technique comes with certain limitations that complicate its application in geomatic/medical images such as the sensitivity to noise. For this reason, an SNN-SELU neural network is used for the denoising phase, guaranteeing the applicability of the proposed methodology.

## 2 Materials and Methods

The proposed methodology consists of a two-stage pipeline. The first stage involves the use of a supervised learning neural network: SNN neural network is modeled to operate the denoising phase of the pipeline. The SNN must identify the noise that if it is present in the brain structure makes it difficult to apply the Watershed algorithm. The second stage, instead, involves the application of the watershed algorithm to the sharpened and processed images, [10], [11].

### 2.1 First Stage: Denoising Process with Self-Normalising Neural Network (SNN)

As known, a typical neural network consists of an input layer and an output layer. The number of neurons in the input layer depends on the specific task and implementation choices, while the size of the output layer varies based on the number of desired output values for the model. The presence of multiple neurons in the output layer can influence the accuracy of the network's predictions. In artificial neurons, inputs are combined with corresponding weights to calculate a weighted sum of the inputs. This weighted sum is then passed through an activation function, which transforms it

into an output. Activation functions play a crucial role in neural networks by mapping input data to output values. This feature is essential for enabling neural networks to learn intricate relationships and patterns within datasets.

Self-normalizing neural networks (SNN), as those used in the study, employ a specific architecture where one of the layers is comprised of neurons that use Scaled Exponential Linear Units (SELU) as activation functions. SNNs are designed to be robust against noise, and they possess the unique characteristic of self-normalization, which means they do not require extensive preprocessing of input data to function effectively. One of the most significant functions in this network is the SELU activation function. SELU, short for Escalated Exponential Linear Unit, plays a crucial role in ensuring stable training and convergence during the learning process. It offers the advantage of unit variance in training errors and convergence toward an average of zero. Notably, SELU is fast and does not require complex initialization methods, making it particularly effective when dealing with noisy training datasets. SELU's properties also promote self-normalization within the network, helping to mitigate the issue of gradient disappearance during training. The combination of Self-Normalizing Neural Networks (SNN) and SELU activation is highly advantageous when designing deep neural networks. This combination ensures that gradients remain stable throughout the training process, enabling networks to learn intricate data representations effectively. In comparison to other activation functions like Rectified Linear Unit (ReLU), SELU is often preferred, especially in convolutional neural networks (CNNs), due to its desirable properties and performance. The SELU (Scaled Exponential Linear Unit) activation function is defined as:

$$SELU(x) = 1.0507 * (x \text{ if } x > 0 \text{ else } (1.6732 * (e^{x-1}))) \quad (1)$$

In this equation:

- `x` is the input to the SELU function.
- `alpha` is a constant with a value of approximately 1.6732.
- `scale` is a constant with a value of approximately 1.0507.
- `e` is the mathematical constant, approximately equal to 2.71828.

The SELU activation function is designed with these specific values of  $\alpha$  and  $\text{scale}$  to help maintain unit variance in training errors and ensure convergence towards an average of zero when weights are correctly initialized. These constants contribute to the effectiveness of SELU in deep neural networks by addressing issues such as vanishing gradients and promoting self-normalization during training, [12], [13].

The neural network training phase is a crucial point of the methodology. The initial step in constructing the training dataset involves subjecting a single "IMi" image to a custom pseudo-random algorithm. This algorithm is crafted using the OpenCV library in conjunction with functions from the Computational Photography and denoising package. The software is applied "n" times to the same image, with variations occurring each time. These variations are introduced randomly and involve altering both the filter parameters provided by the library and the sub-regions of the image where the filter is most relevant. This process generates a set of images labeled as "IMd11 ... IMd1n." Each of these images serves as input for the subsequent segmentation phase, which utilizes OpenCV's Watershed algorithm. This segmentation stage produces a set of images denoted as "IMs11 ... IMs1n." These images contain the segmentation results. To ensure the quality of the produced images, a heuristic algorithm assesses whether they exhibit any segmentation errors. If errors are detected, the corresponding images are discarded, while those without errors are included in the training dataset. It's worth noting that there may be instances of false positives, but the removal process is error-free. The heuristic algorithm rejects images when the generated polygons exhibit an unrealistic number and area. For the elimination of false positives, human intervention is necessary. However, the operator's role is limited to a verification process rather than manual segmentation for network training. To address false positives, a filtering operation was performed, primarily focusing on eliminating glaring errors without employing specific anatomical science-based thresholds. The algorithm employed for creating the training dataset follows a generate-and-test approach. Despite the wide range of possible permutations due to parameter variations, it remains compatible with existing computational resources, [14], [15].

In the final training dataset, each "IMi" image is paired with a set of "IMd11 ... IMd1n" images that have successfully passed the filtering and heuristic assessment stages. The neural network's inference

must be directed towards a function capable of either "flattening" or "highlighting" those pixels whose noisy values could potentially disrupt the effectiveness of the Watershed technique.

In this context:

- "Flattening" refers to the process of reducing or equalizing the intensity values of noisy pixels, making them less likely to interfere with the Watershed segmentation.

- "Highlighting" involves enhancing or emphasizing certain pixels, possibly those representing important features or boundaries, to improve the Watershed segmentation's accuracy.

The neural network's role is to process the input images and produce an output that helps prepare the data for successful segmentation using the Watershed technique. Depending on the specific characteristics of the input images and the noise present, the network should learn to either mitigate noise or enhance critical information, ensuring that the subsequent segmentation process is more robust and accurate.

## 2.2 Second Stage: Watershed Algorithm

The watershed-transformed segmentation technique draws an analogy between a grayscale image and a topographic relief map. In this analogy, each pixel's gray level ( $f(x, y)$ ) is interpreted as its elevation, akin to geographical altitudes, [16], [17]. The process is analogous to how water droplets behave on such a topographic surface, following this process:

*Grayscale as Topographic Relief:* Grayscale images are treated as topographic maps, where pixel intensities signify elevations. Lower-intensity regions correspond to lower altitudes, while higher intensities represent higher elevations.

*Watershed Lines and Collection Basins:* This technique identifies "collection basins" within the image, similar to geographical watersheds. These basins correspond to local minima, essentially the lowest points on the topographic surface. Watershed lines are generated from these local minima, serving as dividing lines that separate different regions or objects within the image.

*Contours and Object Representation:* In the context of image processing, these watershed lines effectively outline the contours of objects within the image. Each object is enclosed within its respective collection basin, with watershed lines acting as boundaries between them.

As known, the Watershed algorithm is based on mathematical principles and processes. The key steps involved the Gradient Calculation computing the gradient of the image to highlight regions of

interest using gradient filters like Sobel or Scharr; Marker Initialization defining markers or seeds to indicate the regions you want to segment within the image; Flood-Fill and Region Growing simulating filling basins from these markers, and regions are grown until they meet at watershed lines; Watershed Lines based on the flooding process and used to segment the image.

In this specific case, the formula for calculating the gradient of an image using the Sobel filter is as follows:

For the gradient in the x-direction:

$$G_x = (I(x + 1, y - 1) + 2 * I(x + 1, y) + I(x + 1, y + 1)) - (I(x - 1, y - 1) + 2 * I(x - 1, y) + I(x - 1, y + 1)) \quad (2)$$

For the gradient in the y-direction:

$$G_y = (I(x - 1, y + 1) + 2 * I(x, y + 1) + I(x + 1, y + 1)) - (I(x - 1, y - 1) + 2 * I(x, y - 1) + I(x + 1, y - 1)) \quad (3)$$

Where:

$I(x, y)$  represents the intensity of the pixel in the original image at position  $(x, y)$ .

$G_x$  is the gradient in the x-direction.

$G_y$  is the gradient in the y-direction.

These formulas calculate the gradient of the image in both the horizontal (x) and vertical (y) directions using the Sobel filter. The resulting gradient can be used to detect changes in image intensity, which often correspond to edges or regions of interest. Subsequent mathematical relationships were applied to define the watershed lines-finding algorithm.

This approach proves effective in detecting the outlines of objects within the image. The Watershed technique excels in detecting object boundaries. This feature makes it valuable in applications that require object recognition and analysis. For example, robotics can be used to identify objects in an environment and make decisions based on such information.

However, the versatility of Watershed extends beyond image processing alone. It is widely used in geomatics for segmenting orographic structures such as hills and mountains, demonstrating its adaptability to a wide range of contexts.

Another advantage is computational speed. When applied efficiently, the Watershed technique can yield rapid results, making it suitable for scenarios that demand fast segmentation. This

feature is particularly useful in emergencies or applications requiring real-time processing.

It should be noted that, despite its numerous advantages, the Watershed technique may present some challenges. Among these, over-segmentation and sensitivity to noise may require additional attention. Consequently, it is often used in combination with other techniques or algorithms to enhance the quality of segmentation.

Below (Figure 1), is an example of the application of the Watershed algorithm's process applied to bidimensional topographic images.

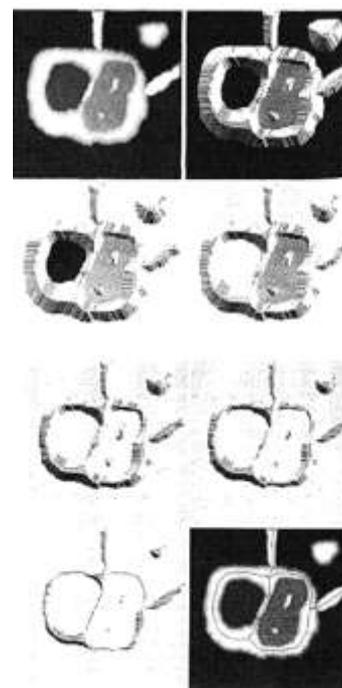


Fig. 1: Watershed process

### 2.3 Information Extraction from Images Dataset (IXI Dataset)

The IXI dataset serves as a valuable resource for brain image segmentation, [18]. It encompasses a vast collection of magnetic resonance imaging (MRI) scans from healthy subjects, spanning different age groups, including both young and elderly individuals. Researchers have developed algorithms for automated brain segmentation based on this extensive dataset. The IXI dataset comprises nearly 600 MR images of healthy individuals, each acquired using a protocol that includes T1, T2, and PD-weighted images, MRA images, and diffusion-weighted images with 15 different directions. Figure 2 shows an example of an image acquired by the IXI dataset.

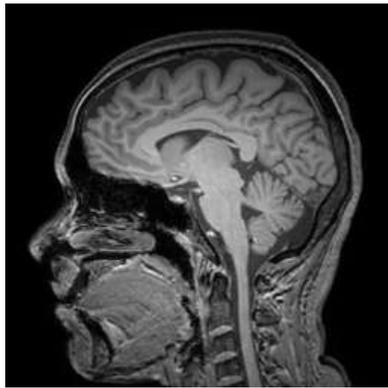


Fig. 2: IXI dataset: IXI263-HH-1684-T1\_70

In the case study, 200 images from the IXI dataset and some images provided by clinical laboratories in the manner prescribed by law were used. The initially generated training set contained approximately 4,200 occurrences, but it was later reduced to about 3,300 occurrences through manual intervention to enhance data quality. The data processing process was carried out independently on several workstations, each equipped with eight 11th-generation Intel Core i7 processors. Each workstation processed one image at a time, and processing was halted if it exceeded a 4-hour computation time. Following the acquisition of the training set, a neural network model based on the Self-Normalizing Neural Network (SNN) with SELU (Scaled Exponential Linear Unit) activation was defined. This type of network is known for its self-normalizing capability, which can contribute to training stability and effectiveness, [19], [20], [21], [22].

Subsequently, the supervised training phase of the neural network was conducted using the prepared dataset.

### 3 Results

After the execution of the training phase, the network's performance within the pipeline framework was assessed, essentially subjecting the methodology to unit testing. To carry out this evaluation, a set of 500 images that were not part of the training set was employed and processed through the pipeline. The main purpose of the proposed method is not so much to accurately identify brain structures, a task entrusted to medical professionals, but rather to provide valuable support during the diagnosis phases where time is limited. This type of segmentation offers significant advantages as it is cost-effective and allows for quick results, which is crucial in emergencies.

However, it is important to emphasize that, despite its speed and efficiency, this technique is capable of segmenting some of the fundamental brain structures, including the optic nerves, pituitary gland, brainstem, and peduncle. In essence, it aims to identify these key regions of the brain to facilitate diagnosis and provide preliminary guidance to medical professionals.

Figure 3 and Figure 4 show the whole pipeline segmentation, denoising stage, and Watershed stage, applied to cerebral images.

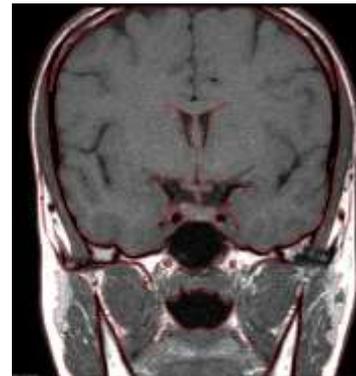


Fig. 3: Complete image processing pipeline, including both the denoising and Watershed stages.

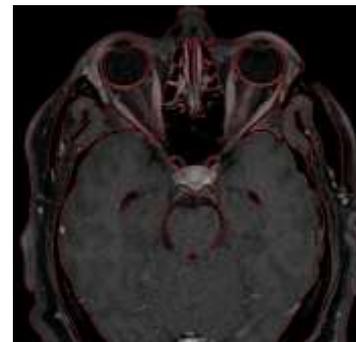


Fig. 4: Example of the whole processing pipeline applied to another cerebral image.

To validate the results obtained, For the sake of a comprehensive analysis, with a primary focus on one of the advantages of the proposed method, which is the reduced image segmentation time compared to traditional methods, we conducted an evaluation using similarity metrics. Specifically, we measured the Intersection over Union (IoU) between manual/ITK-SNAP (software used by experts in the segmentation process) segmentations and the proposed segmentation, [23], [24]. This assessment involved examining different sections of the images to quantify the degree of overlap between the two segmentation approaches.

The results of the similarity analysis, which also took into account false negatives and false positives, revealed an average IoU of 0.93. This value

suggests that the images are generally comparable, indicating strong agreement. However, it's worth noting that there are variations in performance across different regions, with some areas showing excellent results while others may benefit from further improvements.

#### 4 Conclusion

The study explored the application of an approach based on a neural network for image denoising, followed by the utilization of the Watershed algorithm for segmentation in the field of medical images. This methodology was assessed to comprehend its potential and limitations.

The results obtained from this research indicate that the integration of a neural network for denoising with the Watershed algorithm for segmentation offers an intriguing perspective on the management of medical images. However, it is crucial to consider some critical aspects that emerged from the data analysis. One of the main challenges pertains to over-segmentation, a situation where the Watershed algorithm divides image regions into segments that are excessively small or detailed. This can complicate result interpretation and necessitate additional post-processing steps to obtain coherent and clinically meaningful segments.

However, it should be emphasized that, despite this challenge, the methodology presents numerous advantages, including its effectiveness in detecting object contours and its adaptability to various medical applications. Regarding prospects, there are several interesting research directions. The use of larger and more representative training datasets could enhance the denoising neural network's capability.

In conclusion, this research lays the groundwork for further developments in the field of medical image segmentation through the combined use of neural networks and the Watershed algorithm. Despite the challenges, this methodology offers significant opportunities to improve the quality and efficiency of medical image analysis, with a particular focus on image-assisted diagnosis.

#### References:

[1] Zhu, Y., Abdalla, A., Tang, Z., Cen, H. (2022). Improving rice nitrogen stress diagnosis by denoising strips in hyperspectral images via deep learning, *Biosystems Engineering*, Vol. 219, pp. 165-176.

[2] Taher, F., Mahmoud, A., Shalaby, A., & El-Baz, A. (2018, December). A review on the

cerebrovascular segmentation methods. In 2018 *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)* (pp. 359-364). IEEE.

[3] Alirezaie, J., Jernigan, M. E., & Nahmias, C. (1998). Automatic segmentation of cerebral MR images using artificial neural networks. *IEEE Transactions on Nuclear Science*, vol. 45(4), pp. 2174-2182.

[4] Zhu, J., Shi, H., Song, B., Tao, Y., Tan, S., Zhang, T. (2021). Nonlinear process monitoring based on load weighted denoising autoencoder, *Measurement*, Vol. 171, 108782.

[5] Tripathi, S., Sharma, N. (2021). Computer-aided automatic approach for denoising of magnetic resonance images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9:6, pp. 707-716.

[6] Angiulli, G., Barrile, V., & Cacciola, M. (2005). SAR imagery classification using multi-class support vector machines. *Journal of Electromagnetic Waves and Applications*, 19(14), 1865-1872.

[7] Klambauer, G., Unterthiner, T., Mayr, A., Hochreiter, S. (2017). Self-Normalizing Neural Networks. In Proceedings of the 31<sup>st</sup> Conference on Neural Information Processing Systems, Long Beach, CA, USA. *Advances in Neural Information Processing Systems 2017*, 30.

[8] Barrile, V., Candela, G., & Fotia, A. (2019). Point cloud segmentation using image processing techniques for structural analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 187-193.

[9] Yushkevich P. A., Pashchinskiy A., Oguz I., Mohan S., Schmitt J. E., Stein J. M., Zukić D., Vicory J., McCormick M., Yushkevich N., Schwartz N., Gao Y., & Gerig G. (2019). User-Guided Segmentation of Multi-modality Medical Imaging Datasets with ITK-SNAP. *Neuroinform* vol. 17, pp.83-102.

[10] Barrile, V., Cotroneo, F., Genovese, E., & Bilotta, G. (2023). Using Snn Neural Networks Trained with High Resolution Data 2 and Applied to Copernicus SENTINEL-2 Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 48, pp. 27-31.

[11] Barrile, V., Cotroneo, F., Genovese, E., Barrile, E., & Bilotta, G. (2023). An AI Segmenter on Medical Imaging for Geomatics Applications Consisting of a Two-State

- Pipeline, Snn Network and Watershed Algorithm. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 48, pp. 21-26.
- [12] Wang, S., Yang, D.M., Rong, R., Zhan, X., Xiao, G. (2019). Pathology Image Analysis Using Segmentation Deep Learning Algorithms. *Am J Pathol.*, vol. 189(9), pp. 1686-1698.
- [13] Wang, R., Chen, S., Ji, C., Fan, J., Ye Li, Y. (2022). Boundaryaware context neural network for medical image segmentation. *Med. Image Anal.*, vol. 78, pp. 102395.
- [14] Li, P., Jiang, X., Kambhamettu, C., Shatkay, H. (2018). Compound image segmentation of published biomedical figures. *Bioinformatics* 1;34(7):1192-1199.
- [15] Li, H., Chen, C., Fang, S., Zhao, S. (2017). Brain MR image segmentation using NAMS in pseudo-color. *Comput Assist Surg (Abingdon)*, 22(sup1):170-175.
- [16] Khiyal, M. S. H., Khan, A., & Bibi, A. (2009). Modified Watershed Algorithm for Segmentation of 2D Images. *Issues in Informing Science & Information Technology*, 6.
- [17] Acharjya, P. P., Sinha, A., Sarkar, S., Dey, S., & Ghosh, S. (2013). A new approach of watershed algorithm using distance transform applied to image segmentation. *International Journal of Innovative Research in Computer and Communication Engineering*, 1(2), 185-189.
- [18] IXI Dataset – Brain Development. (n.d.). <https://brain-development.org/ixi-dataset/>
- [19] Barrile, V., Cacciola, M., D'Amico, S., Greco, A., Morabito, F. C., & Parrillo, F. (2006). Radial basis function neural networks to foresee aftershocks in seismic sequences related to large earthquakes. *In Neural Information Processing: 13th International Conference, ICONIP 2006, Hong Kong, China, October 3-6, 2006. Proceedings, Part II* 13 (pp. 909-916). Springer Berlin Heidelberg.
- [20] Fu, Y., Lei, Y., Wang, T., Curran W.J., Liu, T., Yang, X. (2020). Deep learning in medical image registration: a review. *Phys Med Biol.* 2020, 22;65(20):20TR01.
- [21] Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T., Yang, X. (2021). A review of deep learning-based methods for medical image multiorgan segmentation. *Phys Med.* Vol. 85, pp.107-122.
- [22] Chen, X., Wang, X., Zhang, K., Fung, K.M., Thai, T.C., Moore, K., Mannel, R.S., Liu, H., Zheng, B., Qiu, Y. (2022). Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal.*, vol. 79, pp.102444.
- [23] Barrile, V., Bilotta, G., Fotia, A., & Bernardo, E. (2020). Road extraction for emergencies from satellite imagery. *In Computational Science and Its Applications–ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part IV* 20 (pp. 767-781). Springer International Publishing.
- [24] Nowozin, S. (2014). Optimal decisions from probabilistic models: the intersection-over-union case. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 548-555).

#### **Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)**

The authors equally contributed to the present research, at all stages from the formulation of the problem to the final findings and solution.

#### **Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself**

No funding was received for conducting this study.

#### **Conflict of Interest**

The authors have no conflicts of interest to declare.

#### **Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

[https://creativecommons.org/licenses/by/4.0/deed.en\\_US](https://creativecommons.org/licenses/by/4.0/deed.en_US)