XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGINEERING

Application of neural networks for the assessment of linking damage zones of geological faults

Arthur Brito Gomes[1], Roberto Quevedo[2], Deane Roehl[1,2], Bruno Raphael Barbosa de Melo Carvalho[3]

[1]*Civil and Environmental Engineering Department, PUC-Rio*
Rua Marquês de São Vicente, 225, 22453-900, Rio de Janeiro, Brazil
arthurbg@aluno.puc-rio.br, droehl@puc-rio.br
[2]*Tecgraf Institute, PUC-Rio*
Rua Marquês de São Vicente, 225, 22453-900, Rio de Janeiro, Brazil
rquevedo@tecgraf.puc-rio.br, deane@tecgraf.puc-rio.br
[3]*Petrobras Research Center*
Av. Horácio Macedo, 950, 21941-915, Rio de Janeiro, Brazil, brcarvalho@petrobras.com.br

Abstract. Damage zones within geological faults are crucial in defining preferential flow paths in reservoirs, impacting oil and gas exploration. Numerical modeling has been used to estimate linking damage zones from fault interactions. However, such methodology requires the knowledge of numerical techniques and is computationally intensive and time-consuming for field-scale studies. This work proposes a deep neural network to predict damage zones resulting from the interaction of two faults, using synthetic data generated through numerical modeling. The dataset was built using key geometric parameters, such as dip angles, overlap, and separation. A modified U-Net with ResNet-50 backbone was trained using a binary classification strategy to identify linked damage zones, employing Dice Loss for optimization. The proposed approach achieved good predictive performance, with an intersection over union (IoU) of 78% and a Dice score of 87% on the test set. The results demonstrate that synthetic models combined with deep learning can efficiently predict fault damage zones. This methodology provides a scalable fast alternative to conventional simulations, enabling systematic studies of fault interactions under controlled geometric conditions.

Keywords: Finite Element Method, Damage Zones, Deep Learning, Geological Faults.

1 Introduction

The definition of damage zones in geological faults is important for oil and gas exploration due to the potential occurrence of preferential flow paths through such zones [1]-[5]. In practice, damage zone widths of single faults are evaluated using exponential laws based on statistical relationships with fault displacement. However, such approaches cannot be applied to linked damage zones resultant from fault interactions. As an alternative, numerical modeling using the Finite Element Method (FEM) combined with elastoplastic constitutive models [6]-[8] has been proposed. In these models, each fault is represented by two superimposed surfaces, with displacements incrementally imposed in opposite directions on each surface. This process generates plastic deformation zones around the faults, which are associated with damage zones. The models have proven effective in identifying linked damage zones through fault interaction simulations. However, the processes of model construction, simulation execution, and result analysis are often time-consuming, potentially delaying decision-making in field operations [9].

The present study proposes the development of a deep neural network to predict damage zones resulting from the interaction of two geological zones. The dataset used for training and testing is based on several numerical models built using a set of geometrical parameters such as overlap, separation, and angle between the faults. This strategy is interesting because it reduces the number of parameters to be sampled while maintain possible retrieve

the global coordinates in the original base. The fault setting as well as the results obtained from the simulations were used as input parameters to train the neural network, identifying the damage zones given by the plastic regions. Then, a binary classification strategy was used to identify damage zones. The preprocessing approach used to predict using the network consists of generating a 256x256 grid points considering a fault arrangement. After a feed forward process, the network estimates the damage zone foreach point of the grid. The preliminary results presented here show the potential of the proposed methodology to quickly assess linking damage zones defined by two geological faults.

2 Dataset

2.1 Numerical Modeling

Each geological fault is represented by two superimposed lines to which are applied opposing displacements simulating fault throw. These displacements induce plastic deformations around the fault associated with the development of damage zones. This approach based on the finite element method and elastoplastic models has been developed and extensively discussed in previous studies [6]-[8]. Figure 1 illustrates the corresponding workflow, including model building, mesh generation and result analysis.

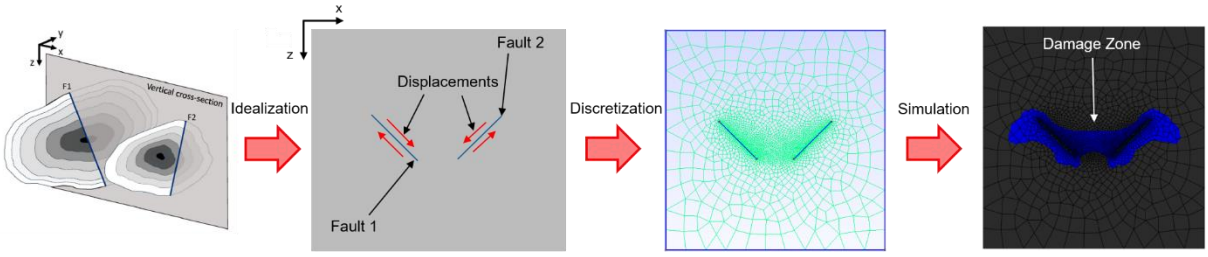
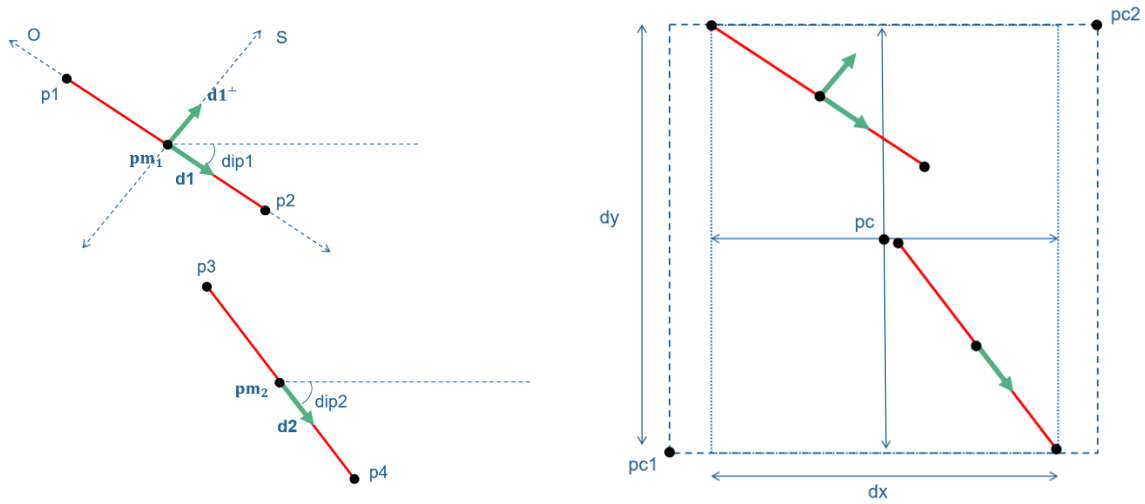


Figure 1. Workflow for numerical modeling.

2.2 Sampling

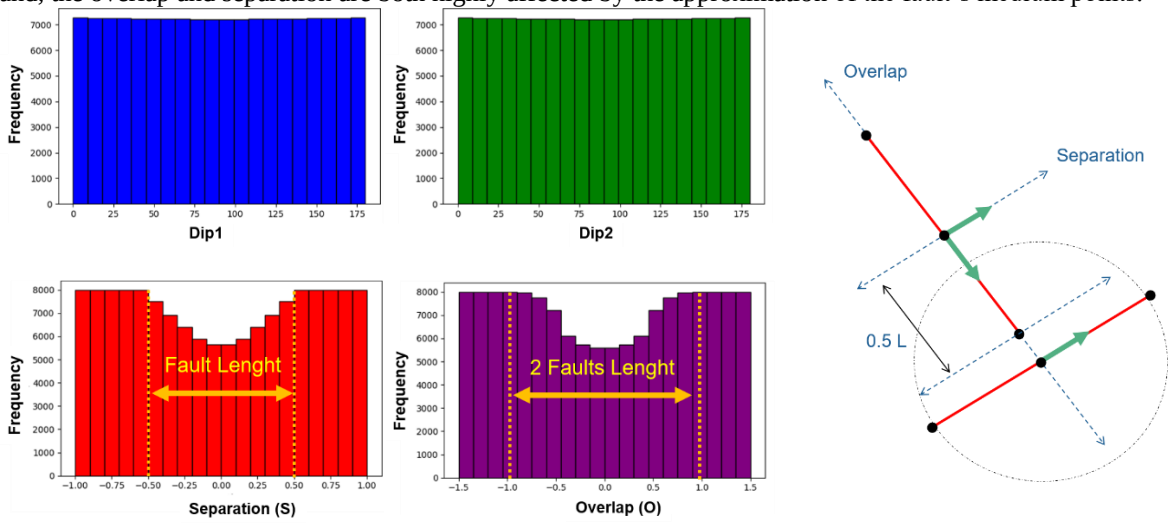
Model geometry was generated using the sketch shown in Figure 2 (a). The main variables are the fault inclinations (dip_1 and dip_2), separation (S) and overlap (O). In addition, some control points and a clearance parameter are used to define limits of the region of interest, as shown in Figure 2 (b).



(a) (b)

Figure 2. Geometry variables of multi fault formulation.

Such geometrical variables are a concise way of defining the region of interest using numerical methods without exhausting the simulation process. Figure 3 (a) highlights the distribution of the modelling space by sampling uniformly the parameters: dip_1, dip_2, O and S. Figure 3 (b) the geometric interpretation of the parameters. The angles range from 0° to 180°, S from -1L to 1L and O from -1.5L to 1.5L. The search space consists of 160 000 models. The resulting distributions of dip_1 and dip_2 are practically uniform. On the other hand, the overlap and separation are both highly affected by the approximation of the fault's medium points.



(a) (b)

Figure 3. Distributions of variables (dip_1, dip_2, S and O).

The dataset was divided in 2 parts, the first one was sampled by a grid with a wide range of overlap and separation (-4L to 4L) to identify the range of parameters that provides an interaction of the damage zones. However, it was observed that their main interaction occurs between -3L and 3L and become increasingly pronounced as close to the fault's midpoint. Then, a Latin hypercube sampling (LHS) was used considering an interval where the interaction between faults is strong (-1L to 1L). The final dataset adding all models included a total of 667 samples randomly split in 60% for training, 20% validation and 20% for testing. Table 1 shows the number of models that contains linked damage zones (DZ) and their percentages in each dataset split.

Table 1. Dataset split and proportions.

Split	DZ Connection	Frequency	Percentage
Train	Yes	246	61.50%
	No	154	38.50%
Validation	Yes	82	61.65%
	No	51	38.35%
Test	Yes	84	62.69%
	No	50	37.31%

3 Training

3.1 Network

A modified version of U-Net [10], [11] was used with ResNet-50 as backbone along the encoding process to improve feature extraction. Only 5 decode channels with sizes 256, 128, 64, 32 and 16 were used with pretrained weights from ImageNet. The number of input channels was 1 due to the binary type of the input. This means that

only the geometry of the faults is used to predict the occurrence of damage zones.

3.2 Loss function

About the minimization process, the Dice Loss [12] function given in Equation (1) was used because of the semantic segmentation property which leads to a region level taxonomy [13]. It is important to predict not only linked damage zones but also isolated faults. Adam [14] was the optimizer with a learning rate equal to 10^{-4}. Two early stopping criteria were adopted: 20 epochs for patience or a difference of loss equal to 10^{-9} between epochs. The maximum number of epochs is 1000 and the batch size 32. The threshold for the sigmoid function was 0.5. Only training and validation splits were used during the training loop.

$$Dice\ Loss = 1 - 2\frac{y_{true} \cap y_{pred}}{y_{true} + y_{pred}} \tag{1}$$

Figure 4 shows the Dice loss minimization over the number of epochs with two important highlights. First, the model is learning relevant patterns as the loss decreases, and second, the best model on validation split occurred at epoch 94, which was stopped by the patience criterion. The model saving criterion played a fundamental role in preventing overfitting. Only the weights corresponding to the lowest validation loss were saved.

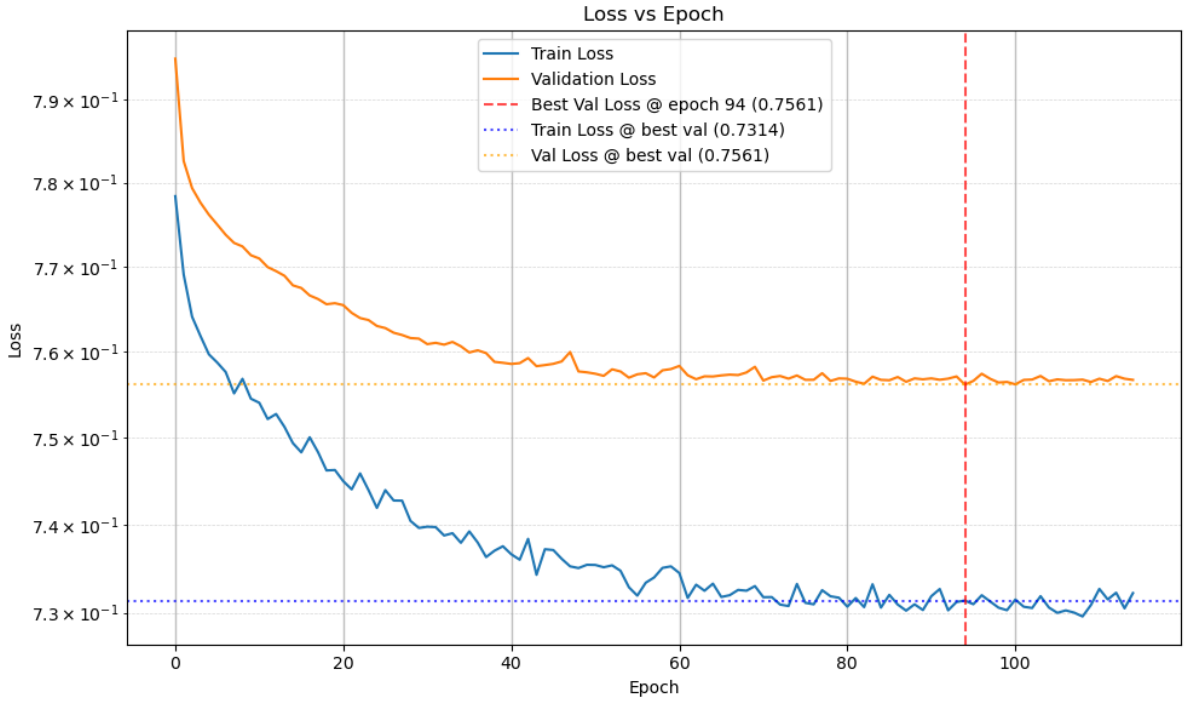


Figure 4. Loss curve over epochs.

4 Results and discussion

The trained model was evaluated using the testing set. The corresponding input data as well as the ground truth and the predictions are shown in Figure 5, which indicate the LHS sampling of faults from -1L to 1L, of O and S. Classification metrics are shown in Table 2, the obtained values agree with the visual results of Figure 5 and Figure 6.

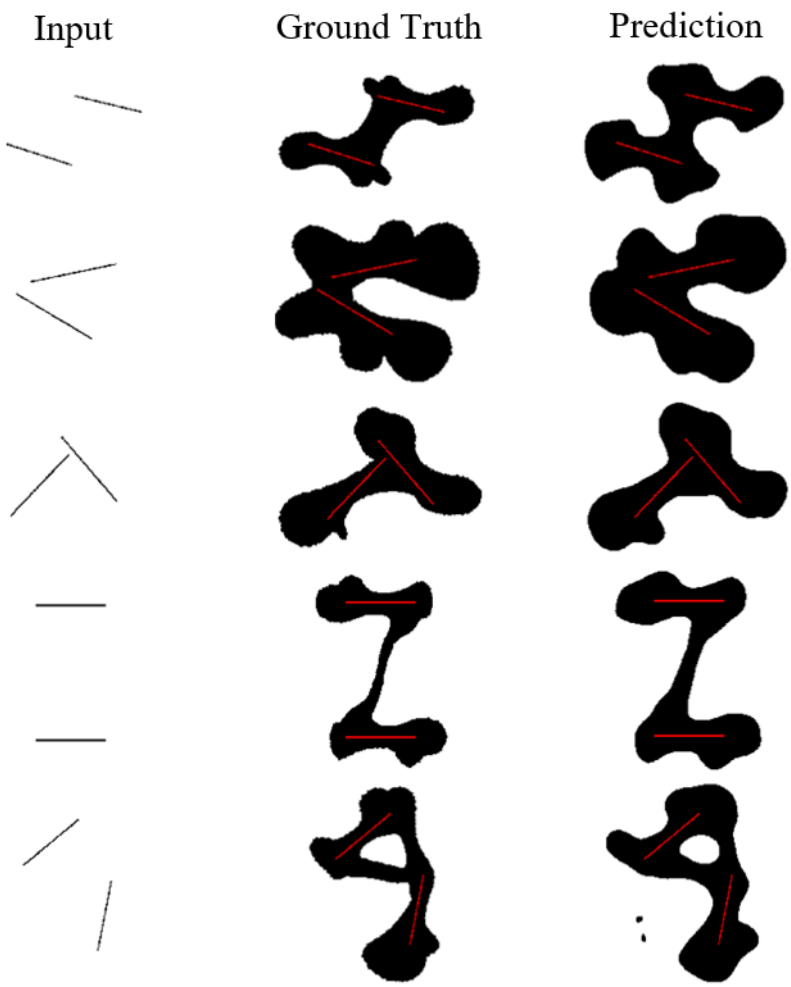


Figure 5. Model predictions on the testing set.

Moreover, high values of O and S are shown in Figure 6. The region of this data belongs to a grid sampling that presented a smaller amount of data. As shown, the region where the fields begin to interact, tending to connect the damage zones, did not present significant changes in some cases, practically maintaining the behavior of isolated faults.

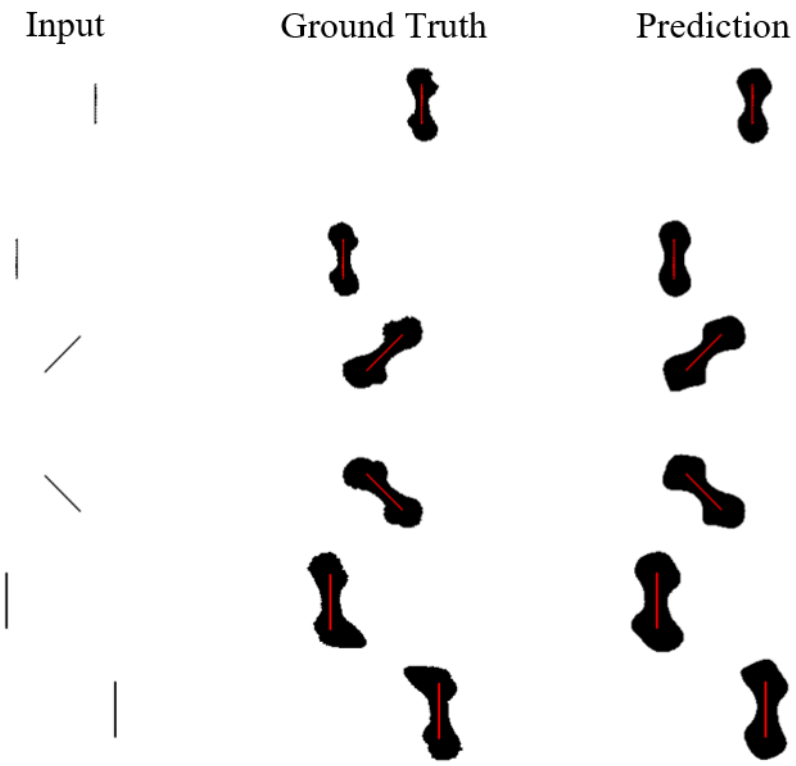


Figure 6. Model predictions to high O and S values.

Table 2. Testing set metrics.

Metric	Mean Values
Accuracy	0.9659
Precision	0.8331
Recall	0.9223
Specificity	0.9745
F1/Dice	0.8725
IoU/Jaccard	0.7787

5 Conclusions

This study presented a methodology for assessing linking damage zones resulting from fault proximity using a neural network trained on the results of several numerical models. Extensive zero padding was applied, which is uncommon in semantic segmentation. The U-Net architecture with a ResNet-50 backbone demonstrated good performance in predicting linking damage zones, as confirmed by both visual inspection and quantitative metrics. The proposed methodology provides an efficient alternative for rapidly generating preliminary predictions of damage zones in scenarios involving interacting geological faults.

Sensitivity analyses indicated that fault throw is the most influential variable in damage zone formation, while variations in rock properties and stress state exert limited influence on the damage zone. Nevertheless, the effect of heterogeneity remains relevant and need further investigations. Building upon these findings, future developments will focus on validating the proposed approach using real geological data provided by Petrobras, enabling the assessment of its applicability under field conditions. Although the present analysis focused on the interaction between two faults, the methodology can be extended to cases involving multiple faults by generating additional synthetic samples representing more complex fault networks. The current implementation was developed for two-dimensional cross-sections, and ongoing work aims to scale the methodology to reservoir-scale 3D models through the combination of multiple 2D sections.

Acknowledgements. This research was carried out in association with the ongoing R&D project registered as ANP nº 21475-9, "GeoBanD 2 – Geomodelagem de zona de dano em falhas geológicas" (PUC-Rio/CENPES/ANP), sponsored by Petrobras. The authors also gratefully acknowledge the support from Carlos Chagas Filho Foundation for Supporting Research in the State of Rio de Janeiro (FAPERJ) Grant E-26/211.766/2021-0, Brazilian National Council for Scientific and Technological Development (CNPq) Grants 141617/2020-9, 420074/2021, and 407388/2022-2, Department of Civil Engineering and Environment, and Tecgraf Institute from PUC-Rio.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] H. Fossen, *Structural Geology*. Cambridge University Press, 2010. doi: 10.1017/CBO9780511777806.

[2] A. Torabi, T. S. S. Ellingsen, M. U. Johannessen, B. Alaei, A. Rotevatn, and D. Chiarella, "Fault zone architecture and its scaling laws: where does the damage zone start and stop?," *Geol. Soc. London, Spec. Publ.*, vol. 496, no. 1, pp. 99–124, Jan. 2020, doi: 10.1144/SP496-2018-151.

[3] S. Mayolle *et al.*, "Scaling of fault damage zones in carbonate rocks," *J. Struct. Geol.*, vol. 124, pp. 35–50, Jul. 2019, doi: 10.1016/j.jsg.2019.03.007.

[4] D. Qu and J. Tveranger, "Incorporation of deformation band fault damage zones in reservoir models," *Am. Assoc. Pet. Geol. Bull.*, vol. 100, no. 03, pp. 423–443, Mar. 2016, doi: 10.1306/12111514166.

[5] J.-H. Choi, P. Edwards, K. Ko, and Y.-S. Kim, "Definition and classification of fault damage zones: A review and a new methodological approach," *Earth-Science Rev.*, vol. 152, pp. 70–87, Jan. 2016, doi: 10.1016/j.earscirev.2015.11.006.

[6] M. Congro, A. S. Zanatta, K. Nunes, R. Quevedo, B. R. B. M. Carvalho, and D. Roehl, "Determination of fault damage zones in sandstone rocks using numerical models and statistical analyses," *Geomech. Energy Environ.*, vol. 36, p. 100495, Dec. 2023, doi: 10.1016/j.gete.2023.100495.

[7] K. Nunes, R. Quevedo, D. Roehl, and B. R. B. M. Carvalho, "Assessment of linking damage zones of geological faults through numerical modeling," *J. Struct. Geol.*, vol. 195, no. February, p. 105381, Jun. 2025, doi: 10.1016/j.jsg.2025.105381.

[8] R. Quevedo, T. J. de Andrade, L. Santos, B. R. B. M. Carvalho, and D. Roehl, "Assessment of fault damage zones in carbonate rocks based on numerical and sensitivity analyses," *Tectonophysics*, vol. 864, no. August, p. 230023, Oct. 2023, doi: 10.1016/j.tecto.2023.230023.

[9] L. F. Santos, R. Quevedo, B. R. B. M. Carvalho, and D. M. Roehl, "Prediction of Fault Damage Zones Using Artificial Neural Networks," in *57th U.S. Rock Mechanics/Geomechanics Symposium*, ARMA, Jun. 2023. doi: 10.56952/ARMA-2023-0473.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *IEEE Access*, vol. 9, pp. 16591–16603, May 2015, doi: 10.1109/ACCESS.2021.3053408.

[11] E. Manos, C. Witharana, M. R. Udawalpola, A. Hasan, and A. K. Liljedahl, "Convolutional Neural Networks for Automated Built Infrastructure Detection in the Arctic Using Sub-Meter Spatial Resolution Satellite Imagery," *Remote Sens.*, vol. 14, no. 11, p. 2719, Jun. 2022, doi: 10.3390/rs14112719.

[12] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10553 LNCS, pp. 240–248, Jul. 2017, doi: 10.1007/978-3-319-67558-9_28.

[13] R. Azad *et al.*, "Loss Functions in the Era of Semantic Segmentation: A Survey and Outlook," pp. 1–24, Dec. 2023, [Online]. Available: http://arxiv.org/abs/2312.05391

[14] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, Jan. 2017, [Online]. Available: http://arxiv.org/abs/1412.6980