A Novel Approach to Pipeline Feature Identification Using Semantic Segmentation

Amir Behbahanian, Adrian Belanger, Robert Coelman,
Paul Dalfonso, Ron Lundstrom
T.D. Williamson

Pipeline Pigging and Integrity Management Conference

February 12-16, 2024

Organized by
Clarion Technical Conferences

https://doi.org/10.52202/072781-0006

https://doi.org/10.52202/072781-0006

Abstract

Machine learning tools have been used for over a decade to process the large amounts of in-line Inspection (ILI) data and to report accurate sizing of metal loss features. With the explosion of software solutions that came with the advancement of graphic processing units, (GPU) power and memory available to the commercial market, the use of machine learning in processing magnetic flux leakage (MFL) tools for corrosion sizing are now standard in the industry. An aspect of automation that does not get as much attention is the identification of features such as fixtures. Traditionally this task has been done either manually or with a semi-automated process based on signal pattern recognition by utilizing operator provided survey data. As fixtures are commonly used as a reference point when performing dig verifications, fast and effective methods for locating them are helpful.

This paper will present an approach using a machine learning method called semantic segmentation. There is an enormous amount of information contained in an MFL survey that has an image like structure with magnetic measurements sampled both axially and circumferentially on a grid. The powerful technique of semantic image segmentation, which is used in applications like autonomous driving, medical imaging, urban planning, manufacturing, robotics, etc., is ideal for analyzing this data.

Four different models will be presented to demonstrate how a combination of them can perform the task of identifying and classifying pipeline features. This work will lay the foundation of not only automating a semi-manual task but training models to focus on features of concern to the operator. Due to concerns about aging infrastructure, identifying certain types of fixtures can help with integrity threat management. The use of these models has the potential of identifying specific features that have known integrity concerns and facilitate a pipeline owner's remediation plan.

Keywords: Semantic Segmentation, Magnetic Flux Leakage, MFL, FPN, Feature Pyramid Networks.

Introduction

Since the 1960s, the examination and maintenance of pipeline networks have relied on the Magnetic Flux Leakage (MFL) phenomenon. [1] Understanding and relating the MFL signal traits to pipeline defects and categorizing pipeline components have posed significant challenges in the inspection sector. This issue is challenging, existing within a multi-dimensional framework influenced by factors such as pipeline materials, manufacturing processes, internal conditions created by the product being transported through the pipeline and the non-unique inversion problem of mapping MFL to metal loss geometry. Moreover, data collected during inspections may exhibit a low signal-to-noise ratio due to the dynamics of the varying distance of sensors to pipe the pipe surface, magnetic permeability variations due to pipe metallurgy and manufacture, and changes in the tool velocity as it performs its inspection. To tackle these hurdles, the industry has combined experimental[3], signal processing[4]-[6], and numerical techniques[7]. With the universal approximation theorem enabling Neural

https://doi.org/10.52202/072781-0006

Networks (NN) to handle complex functions [8], it is anticipated that neural networks will surpass traditional methods in prediction accuracy, given the complexity of this problem.

Similarly, both academia and industry frequently integrate NNs into their methodologies. These efforts entail replicating finite element solutions to Maxwell's equations, which constitute the forward process of characterizing the MFL generated by metal loss geometries[9][10] and developing a depth prediction by parameterizing signals into tabular forms[11][12]. These endeavors to resolve Maxwell's equations and compare signals from an inspection with a database of solutions represent a computationally intensive approach. Additionally, the matching process of MFL signals to depths poses challenges, given the nonlinear nature of Maxwell's equations, especially in the presence of nonlinear materials like pipe steel with various manufactures. The process of summarizing signals into tabular formats due to signal sampling leads to a reduction in information compared to examining signals comprehensively [13]. Analyzing MFL signals through convolutions while maintaining signal integrity, facilitates the utilization of more intricate Convolutional Neural Networks (CNNs) capable of generalizing to complex defect interactions. This is made possible by accessing pixel- level information [14]–[16].

Semantic Segmentation, a well-established technology in various fields like medical imaging[17], autonomous driving[18], and surveillance[19], involves image classification at a pixel-level. The journey of image classification models commenced in 1998 with the LeNet model [20]. However, the inclusion of Multi Perceptron Layers (MLP) at the end of these models posed limitations on achieving pixel-level classification. Subsequent advancements extended the LeNet model, leading to iterations like AlexNet [21], VGGNet[22], and GoogleNet [23]. The advent of Fully Convolutional Networks (FCN) marked a milestone by enabling pixel classification, thereby enhancing object detection boundaries.

Various strategies emerged to refine classification quality, focusing on enhancing information flow and preservation within the encoder section [24][25]. Moreover, techniques were devised to maintain precise object boundaries by establishing path information between the encoder and decoder within the architecture. This conceptualization gave rise to diverse architectures like UNet[26], Linknet[27], FPN [28], and PSPNet[29] among others. These architectures have demonstrated success in extracting and generalizing information from datasets across diverse environments, such as cities and various weather conditions[19], providing evidence for their inherent inductive bias[30].

Considering the gaps in the current approaches to analyzing MFL signals and the promising generalizability of results from encoder-decoder architectures, we employ the use of the approach to analyze MFL data. As an initial attempt of incorporating the models into our workflow, we solve the problem of detecting pipeline components. This step is an essential piece of pipeline integrity analysis that enables the correlation of extracted analyses with observations in the field. In our current work, we elucidate the methods and approaches we used for our previously developed and published work [31]. We also provide data curation approaches to make the training process feasible and provide

https://doi.org/10.52202/072781-0006

high performance. This work further establishes the potential of encoder-decoder networks to analyze MFL data as well as their capability to capture the complexities of the physics, of adding different inspection conditions, and to provide reliable industry accepted results.

Methodology

Each pipeline inspection, called a run, encompasses the examination of multiple connected pipe joints comprising a section of the inspected pipeline. The analysis presented here draws data from 209 runs managed by various pipeline operators in North America, ensuring a diversified dataset that accurately represents pipeline configurations in North America. For this analysis, we isolated and stored MFL data separately for each joint within the pipeline, focusing solely on individual pipe joint features. This decision stemmed from the recognition that pipeline features spanning multiple joints exhibit limited diversity, falling within a narrow range of categories.

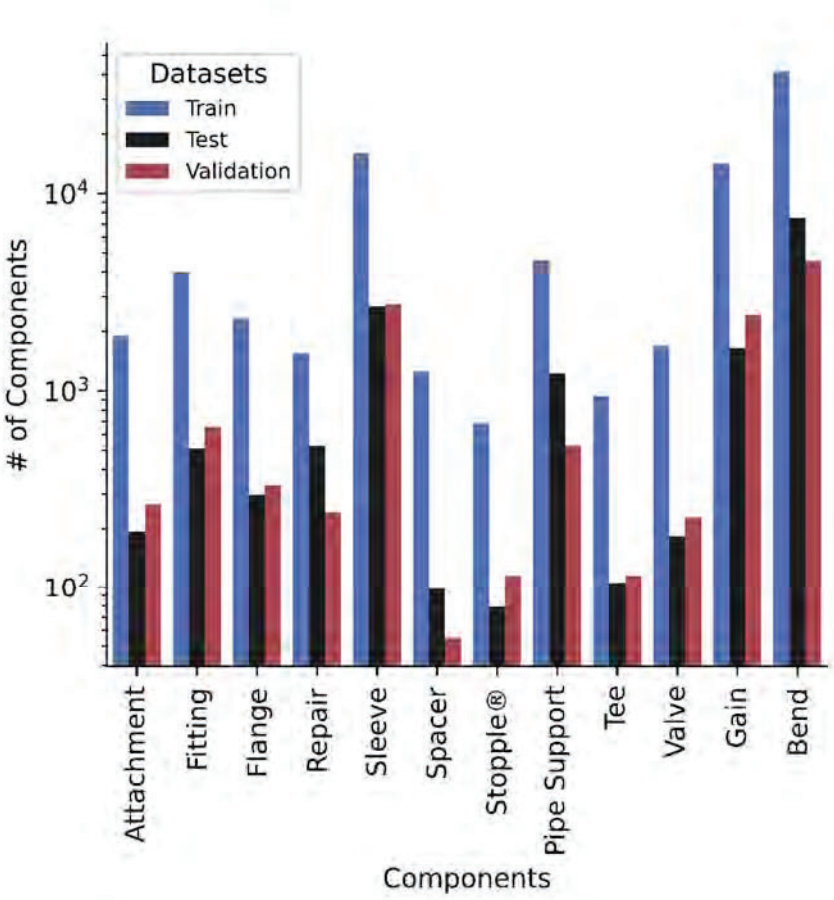

Figure 1 : Breakdown of the number of training, testing and validation sets for each pipeline component.

https://doi.org/10.52202/072781-0006

With a dataset comprising 115 thousand pipe joints, we performed random selections for training, validation and test sets, represented in Figure 1, with a distribution ratio of 80%, 10%, and 10% respectively. The components in the training set are used to determine the weights in the model by minimizing a cost function, which is a function that effectively measures the error between the model prediction and the actual value. The components in the validation set are used to evaluate the model. Every model has hyperparameters which can be tuned, and the validation set is held back from training the model in order to monitor the performance of the model unbiasedly during training time. Finally, the model is evaluated against a test set that contains data it has never seen either in training or in validation.

The first stage of the model development is to determine the hyperparameters and uses only sampling of the total data divided up into training, validation and test sets in the relative proportions previously stated. This stage determines the loss function and model optimization hyperparameters that will be used in the next stage. The next stage uses all the data divided up in the 80% - 10% - 10% proportion for the training, validation and test sets respectively. The model is solved in stages called epochs where the model fitted to the training data is compared against the validation data and then retrained. At the end of all the epochs the results are compared against the test data. The performance against this test set is the best predictor on how the model will perform on future runs. on operational to act

Across all three datasets, in addition to background, twelve primary classes were identified The background class encapsulates crucial details about pipeline defects, presenting a scale smaller than the primary components. Consequently, minimizing pipeline data size reduces the background information to an undetectable scale, ensuring uniformity across the background as a singular class.

The dataset's class distribution (Figure 1) reveals a significant imbalance among the classes, emphasizing the criticality of selecting an appropriate loss function. Moreover, aligning with business objectives, balancing the extra work of false positives with the risk of false negatives, adds another layer of consideration while evaluating loss functions. Given these parameters and the semantic segmentation nature of the problem, our focus narrowed down to three specific loss functions: Dice [32], Lovasz-Softmax [33] and Focal Tversky [34].

The Dice Loss (DL) is a commonly utilized loss function that measures image similarity. However, its equation exposes the potential for gradient vanishing when dealing with small values of $\boldsymbol{y}$ (actual) and $\hat{\boldsymbol{y}}$ (predicted):

$$\boldsymbol{DL(y,\hat{y}) = 1 - \frac{2y\hat{y}}{y + \hat{y} + 1}}$$

On the other hand, the Lovasz-Softmax loss optimizes the Jaccard index [35], the ratio of the intersection of two sets divided by their union, exhibiting enhanced performance in terms of intersection over union. Further research indicates its efficacy in accurately capturing object boundaries [36].

https://doi.org/10.52202/072781-0006

Meanwhile, the Focal Tversky loss (FTL), derived from the Dice loss, specifically tackles data imbalance issues, achieving a balanced precision and recall.

$$\boldsymbol{FTL(y,\hat{y}) = \left(1 - \frac{y}{y + \alpha\hat{y} + \beta y + 1}\right)^{\gamma}}$$

Leveraging the Tversky index, which allows model tuning to favor either false positives or false negatives, FTL adds another parameter γ which allow tuning for less accurate predictions that have misclassified data points. The α and β coefficients allow tuning against false negatives and false positives, respectively, which aligns well with tailoring the model to fulfill specific business objectives.

Modern fully convolutional neural networks tailored for semantic segmentation typically consist of an encoder and a decoder. The encoder's primary role involves feature extraction from images while reducing the feature space dimensions through manifold learning techniques which are just techniques that map data in higher dimensions to a lower dimension. The architectural variations in the encoder section of these models aim to optimize information pathways, preserving maximum information and establishing a diverse feature space to enhance prediction accuracy.

Conversely, the decoder component undertakes the task of up-sampling the feature space to match the original image size, contributing to the final prediction. Designing a complete encoder-decoder model also involves the crucial aspect of retaining intricate details during the image up-sampling process. This necessity leads to various methodologies for transferring information between the encoder and the decoder.

In our quest to identify the most effective architecture, we opted for FPN and Unet due to their demonstrated performance in other applications [37][38]. FPN is a top-down architecture that creates features on each level of the decoder and independently makes predictions. Unet is a fully convolutional network like FPN, but it concatenates feature maps at the same spatial scale and predictions are only made on the final upsampled feature map. Combining these architectures, we integrated three distinct encoders—Xception [39], ResNeXt [40], and DenseNet [25]—in order to leverage their strengths in enhancing the segmentation performance of our model.

Xception is an encoder that produces depthwise separable convolutional layers. It operates on the assumption that all the convolutions are entirely separable and creates a combined stack of features on a layer that it passes on to the next layer. The ResNeXt encoder is designed to transform points sharing the same topology and project them onto a lower dimension. Instead of adding multiple layers, ResNeXt preserves information via a twostep process that adds the identity data point to its transformation. This encoder performs gradient propagation as the error can be back propagated along multiple paths. Finally, the DenseNet encoder is an extension of ResNeXt where instead of just connecting an encoding layer to the layer after it, it connects to all layers after it allowing information to be persevered in the downsampled layers.

https://doi.org/10.52202/072781-0006

After building the model, we conducted optimization runs separately from the main training procedure, utilizing a random subselection of our data, to fine-tune the loss function and parameters specifically for the Focal Tversky loss. Our findings indicated superior performance of the Focal Tversky loss in comparison to Softmax-Lovasz and Dice. Additionally, we optimized the values of α and β, governing the model's predictions by penalizing false positives or false negatives. The optimized values for α and β were determined thus minimizing false negatives and aligning with our business objectives [31].

Results

Our training process involved using the Focal Tversky loss on four models, conducted on two NVIDIA Tesla M60s, managed through the PyTorch Lightning module [41] across twenty epochs. Results indicated that the DenseNet encoder with the FPN architecture emerged as the top-performing model. However, our experiments suggested that a weighted ensemble model, utilizing logits predictions, outperformed all other models in metrics.

Attachment values are reported as zero due to most instances being classified as either Fittings or Gains. This occurs because Attachments bear similarity to these categories within the MFL dataset, often necessitating additional data sources for accurate identification. Notably, some classes encompass diverse pipeline components, such as Gains and Repairs, with signals that might appear random to the model due to their varied definitions and appearances. Consequently, these classes exhibit lower precision and recall rates [31].

The majority of classes maintain an acceptable recall, aligning with our project's business objectives. Lower precision in certain classes, like fittings, may partly result from missing components in the ground truth detected by the model (Fig. 2 rows 5 & 8). Furthermore, the model's classification and boxing performance have streamlined our data analysis team's reporting process. Their tasks now focus on editing the model's results rather than detecting, classifying, and boxing components. Remarkably, the model even identified objects absent in our ground truth, signifying an enhancement in detection quality (Fig. 3 row 8).

https://doi.org/10.52202/072781-0006

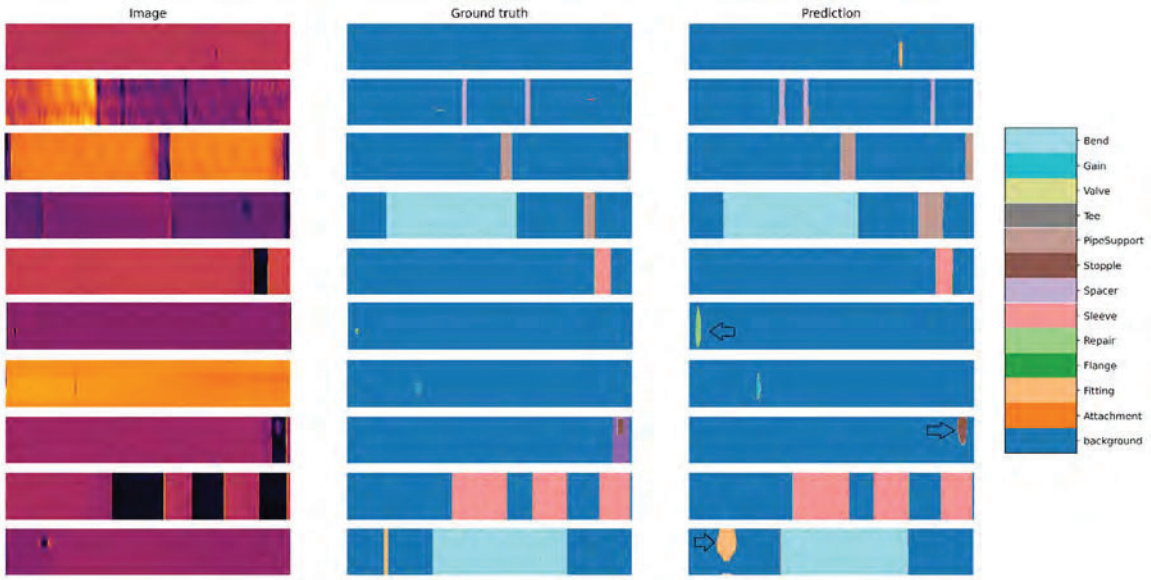

Figure 2 : Comparison of the MFL data (left) to the annotated features (middle) and predicted features (right)

This approach not only reduces the reporting time but also enhances the quality of our data analysis team's work by replacing repetitive tasks with an editing procedure.

Conclusion

Our efforts culminated in the development of a network capable of classifying and boxing pipeline components using MFL (Magnetic Flux Leakage) data. To create these models, we initially selected a loss function from three candidates using a subset of the data. Subsequently, we evaluated four combinations of encoders and architectures, namely DenseNet-FPN, ResNeXt-FPN, ResNeXt-UNet, and Xception-FPN, using our complete training dataset. Among these models, DenseNet-FPN demonstrated the best performance. Further analysis revealed that the most exceptional performance was achieved through a hybrid model, a blend of all four models. This hybrid approach involves aggregating logits from each model and applying a weighted average, resulting in superior performance.

Consequently, we successfully achieved our objective of classifying and boxing pipeline components while significantly reducing the operational and reporting time of our data analysis team. Their workload transitioned from manual detection, classification, and boxing to editing the output generated by the developed model. Moreover, by eliminating the more tedious aspects of their work, we've notably enhanced the accuracy of our data analysis efforts. The model's capability to detect and classify objects missing in the mask further supports the argument for improving the quality of data analysis work.

https://doi.org/10.52202/072781-0006

This project holds promise in enabling our data analysis team to effectively manage pipeline system integrity, optimize maintenance spending, and mitigate the risks of potential incidents. Moving forward, our plan involves ongoing refinement of the model to enhance its performance and extend the utilization of similar models for other pipeline inspection tasks.

Acknowledgments

We express our sincere gratitude to Cory Estes and Jed Ludlow, who provided invaluable support to this project in their roles as the TDW Senior Director of Engineering and Chief Engineer. Additionally, we extend our heartfelt appreciation to Dane Burden, Shan Li, David Sunwall, Matt Romney, Mike Kirkwood, and Semyon Bokhankevich for their insightful comments and valuable inputs, which greatly contributed to the success of this endeavour.

References

1. K. Mandal and D. Atherton, "A study of magnetic flux-leakage signals," *Journal of Physics D: Applied Physics*, vol. 31, no. 22, p. 3211, 1998.
2. M. Afzal, S. Udpa, L. Udpa, and W. Lord, "Rejection of seamless pipe noise in magnetic flux leakage data obtained from gas pipeline inspection," *AIP Conference Proceedings*, vol. 509, no. 1. American Institute of Physics, 2000, pp. 1589–1596.
3. X. Peng, U. Anyaoha, Z. Liu, and K. Tsukada, "Analysis of magnetic-flux leakage (MFL) data for pipeline corrosion assessment," *IEEE Transactions on Magnetics*, vol. 56, no. 6, pp. 1–15, 2020.
4. M. Afzal and S. Udpa, "Advanced signal processing of magnetic flux leakage data obtained from seamless gas pipeline," *NDT & E International*, vol. 35, no. 7, pp. 449–457, 2002
5. M. Afzal, J.-J. Kim, S. Udpa, L. Udpa, and W. Lord, "Enhancement and detection of mechanical damage MFL signals from gas pipeline inspection," *Review of Progress in Quantitative Nondestructive Evaluation*, vol 18A–18B, pp. 805–812, 1999
6. M. A. K. Afzal, *Wavelet based multiresolution zero-crossing representations*. Iowa State University, 2000
7. P. Ramuhalli, "MFL signal inversion using the finite element network," *AIP Conference Proceedings*, vol. 975, no. 1. American Institute of Physics, 2008, pp. 633–640
8. L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, "Learning nonlinear operators via deeponet based on the universal approximation theorem of operators," *Nature machine intelligence*, vol. 3, no. 3, pp. 218–229, 2021.
9. P. Ramuhalli, *Neural network based iterative algorithms for solving electromagnetic NDE inverse problems*. Iowa State University, 2002
10. M. Le, C.-T. Pham, and J. Lee, "Deep neural network for simulation of magnetic flux leakage testing," *Measurement*, vol. 170, p. 108726, 2021

https://doi.org/10.52202/072781-0006

11. Khodayari-Rostamabad, J. P. Reilly, N. K. Nikolova, J. R. Hare, and S. Pasha, "Machine learning techniques for the analysis of magnetic flux leakage images in pipeline inspection," *IEEE Transactions on magnetics*, vol. 45, no. 8, pp. 3073–3084, 2009
12. J. Feng, Y. Yao, S. Lu, and Y. Liu, "Domain knowledge based deep-broad learning framework for fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 4, pp. 3454–3464, 2020
13. Belanger, D. Burden, and P. Dalfonso, "Not all data is good data: The challenges of using machine learning with ILI," IPC2022-86934, *Preceedings of the 2022 International Pipeline Conference*, vol. 86571, American Society of Mechanical Engineers, 2022, V002T03A023.
14. W. S. Rosenthal, S. Westwood, and K. Denslow, "Pipeline defect detection and fine scale reconstruction from 3-d MFL signal analysis using object detection and physics-constrained machine learning," IPC2022-87313, *Preceedings of the 2022 International Pipeline Conference*, vol. 86571, American Society of Mechanical Engineers, 2022, V002T03A060.
15. K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian, "Centernet: Keypoint triplets for object detection," *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6569–6578.
16. Z. Li, *Deep Learning Techniques for Magnetic Flux Leakage Inspection with Uncertainty Quantification.* Michigan State University, 2019.
17. S. Asgari Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, and G. Hamarneh, "Deep semantic segmentation of natural and medical images: a review," *Artificial Intelligence Review*, vol. 54, pp. 137–178, 2021.
18. D. Feng, C. Haase-Schütz, L. Rosenbaum, H. Hertlein, C. Glaeser, F. Timm, W. Wiesbeck, and K. Dietmayer, "Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges," *IEEE transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1341–1360, 2020
19. Y. Mo, Y. Wu, X. Yang, F. Liu, and Y. Liao, "Review the state-of-the-art technologies of semantic segmentation based on deep learning," *Neurocomputing*, vol. 493, pp. 626–646, 2022.
20. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
21. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017
22. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, arXiv:1409.1556, 2014.
23. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions (2014)," *arXiv preprint*, arXiv:1409.4842, vol. 10, 2014.

https://doi.org/10.52202/072781-0006

24. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778
25. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
26. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer, 2015, pp. 234–241.
27. Chaurasia and E. Culurciello, "Linknet: Exploiting encoder representations for efficient semantic segmentation," *2017 IEEE visual communications and image processing (VCIP)*. IEEE, 2017, pp. 1–4
28. T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125
29. H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.
30. T. Mansour et al., "Deep neural networks are lazy: on the inductive bias of deep learning," Ph.D. dissertation, Massachusetts Institute of Technology, 2019
31. Behbahanian, Amir; Lundstrom, Ron; Belanger, Adrian; Dalfonso, Paul; Coleman, Robert, 'PIPENet: A Semantic Segmentation Approach to Pipeline Component Detection from Magnetic Flux Leakage Readings", *International Conference on Machine Learning and Applications (ICMLA),* 2023, IEEE.
32. C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop*, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3. Springer, 2017, pp. 240–248
33. M. Berman, A. Rannen Triki, and M. B. Blaschko, "The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4413–4421
34. N. Abraham and N. M. Khan, "A novel focal Tversky loss function with improved attention u-net for lesion segmentation," 2019 *IEEE 16th international symposium on biomedical imaging* (ISBI 2019). IEEE, 2019, pp. 683–687
35. S. V. des Sciences Naturelles, *Bulletin de la Société vaudoise des sciences naturelles*. Librairie F. Rouge & Cie, 1864, vol. 7.

https://doi.org/10.52202/072781-0006

36. L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818
37. J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint* arXiv:1804.02767, 2018
38. X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
39. F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258
40. S. Xie, R. Girshick, P. Doll´ar, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1492–1500
41. W. Falcon et al., "Pytorch lightning," *GitHub*. Note: https://github.com/PyTorchLightning/pytorch-lightning, vol. 3, 2019.

https://doi.org/10.52202/072781-0006