Oil Spill Detection Using Deep Learning
A Comparison of U-Net, DeepLabV3 & PSPNet

Harshit1, Devroop Das2, Gavish Kumar3 and G.Saranya 4

1 Department of Networking and Communication, School of Computing College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India
2 Department of Networking and Communication, School of Computing College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India
3 Department of Networking and Communication, School of Computing College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India
4 Department of Networking and Communication, School of Computing College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu 603203, India

hs6061@srmist.edu.in , dd2239@srmist.edu.in, gs0652@srmist.edu.in and saranyag3@srmist.edu.in
Corresponding Author : G.Saranya

Abstract. Oil spillage is a serious environmental risk, resulting in damage to marine life and human commerce. Timely and precise identification of oil spillage is important for effective response systems. This research evaluates the performance of recent deep learning architectures—U-Net, DeepLabV3, and PSPNet—for oil spill detection using satellite imagery. Comparison of the models was done to determine the respective strengths and limitations of the models to detect areas of oil contamination. The results show that DeepLabV3 was the most accurate, followed by U-Net, and PSPNet. The research highlights the capability of deep learning models for improved detection of oil spills and shows the application of artificial intelligence in environmental monitoring and disaster relief initiatives.

Keywords: Oil Spill Detection, Deep Learning, Remote Sensing, U-Net, DeepLabV3, PSPNet

1. Introduction

Oil spills present a serious hazard to the environment, causing large-scale ecological and economic losses to marine and coastal ecosystems and, thus, disturbing both human activities and biodiversity. Such events normally occur as a consequence of accidents during oil extraction, transportation, or storage operations and, in turn, potentially produce long-term ecological and economic impacts. Large previous events of major oil spills, like the 2010 Deepwater Horizon incident, have shown the long-term impacts of large oil release on the environment, resulting in large-scale contamination and expensive remediation. Early and effective detection of oil spills is essential in an effort to reduce damage and effectively coordinate response activities.

Oil spill detection effectively is critical to the success of mitigation and environmental conservation. Past oil spill detection practices rely on remote sensing technology such as SAR and optical images, but not all the techniques are precise when the weather and the ocean surface is constantly changing. Advances in deep learning have enabled more powerful oil spill detection with CNNs used in classifying and segmenting oil spills from satellite images more accurately. Studies such as Guo et al. (2020) [6] have shown how CNN based models can identify oil spills from polarimetric SAR data, telling the future in preventing oil spills.

Deep learning has shown potential when it comes to image segmentation. The use of CNNs have changed the way satellite images are processed and oil spills are detected. Compared to conventional techniques, CNN based models have shown learning capability for spatial features on their own and respond differently to different spill-types. The performance of the model depends on the architecture and its ability to identify sophisticated spatial features.

This study explores the performance of the three deep learning model which are U-Net, DeeplabV3, and PSPNet to find oil spills through the use of satellite imagery. Each model tackles specific problems of image segmentation in their own way. U-Net has an encoder-decoder structure, which makes it give faster results and high precision. DeeplabV3 uses the atrous spatial pyramid pooling (ASPP) to process multiple scale contextual information. Similarly, PSPNet uses a pyramid pooling module which integrates both global and local features, resulting in better accuracy.

To compare all these models to each other, we have used performance metrics such as IoU i.e Intersection over Union, F1-score and accuracy. We have trained the models over a publicly available dataset from Kaggle.

Recent developments in this field have shown how models such as DeeplabV3+ have performed extremely well in segmentation tasks when compared to other models. But it requires high computational limits as seen in Krestenitis et al. [2].

2. Literature Survey

Solberg, A., Brekke, C., and Husoy, P. "Oil Spill Detection in Radarsat and Envisat SAR Images." IEEE Transactions on Geoscience and Remote Sensing 2007 [1]
Solberg, Brekke, and Husoy (2007) used RADARSAT and ENVISAT imagery, which was highly successful in identifying spills. The research showed the model's excellence in adverse weather conditions.

Krestenitis, M., Oikonomou, M., and Kalogiros, J. "A Benchmark Dataset for Oil Spill Detection Using Deep Learning." Remote Sensing 2019. [2]
Krestenitis et al. (2019) dataset highlighted the importance of using well annotated data in training of such deep learning models. The dataset acts as a primary resource in correct detection and prevention of spills.

Abdimanap, G., Omarov, Y., and Abdallah, A. "ROSID: A Remote Sensing Dataset for Onshore Oil Spill Detection." Environmental Monitoring and Assessment 2023. [3]
Abdimanap, Omarov, and Abdallah (2023) used the ROSID dataset, which had labelled data making it better for deep learning models to be trained over. This dataset solved the problem of detection using the technique of remote sensing.

Mahmoud, H., Zhang, J., and Kim, T. "A Dual Attention Model for SAR-Based Oil Spill Detection Using Deep Learning." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2022. [4]
Mahmoud et al. (2022) used a dual attention mechanism which helped the model focus on particular areas of spills in the SAR imagery, which led to higher accuracy by reducing background noise and false predictions.

Li, Z., Wang, S., and Liu, Q. "Object-Oriented Oil Spill Detection from SAR Images Using Adaptive Local Thresholding and Fuzzy Classification. [5]
Li, Wang, and Liu (2018) integrated the use of adaptive local thresholding and fuzzy classification. This approach was particularly successful because it emphasized on spatial characteristics and texture of the oil-spilled areas

Guo, H., Xu, Y., and Chen, L. "CNN-Based Oil Spill Identification Using Polarimetric SAR Data." [6]
Guo, Xu, and Chen (2020) used CNNs, which showed how polarimetric SAR data could improve spill detection performance and performed significantly better than conventional methods.

Shaban, M., Al-Ali, M., and Habib, A. "A Two-Stage Deep Learning Approach for Oil Spill Detection in Highly Unbalanced SAR Data." ISPRS Journal of Photogrammetry and Remote Sensing 2021. [7]
Shaban, Al-Ali, and Habib (2021) came up with a two-stage deep learning method which came out to be useful in the case of imbalance SAR data. This class imbalance problem was solved by this method, resulting in better accuracy and detection.

Fan, Y., Zhao, P., and Sun, W. "Feature Merge Network for Oil Spill Detection Using Synthetic Aperture Radar Imagery." Journal of Environmental Management 2022. [8]
Fan, Zhao, and Sun (2022) came up with the idea of merging multi-source features from SAR images. Their model worked better in diverse environments and spill conditions.

Rousso, R., Patel, K., and Meena, S. "Enhancing Oil Spill Detection with Image Filtering and Deep Learning Models." Geosciences 2023. [9]
Rousso, Patel, and Meena (2023) integrated image filtering methods. This resulted in removal of irrelevant noise, making the model concentrate on significant oil spills improving the overall performance.

Ma, L., Singh, V., and Zhao, T. "Dual-Polarimetric Sentinel-1 SAR for Oil Spill Detection Using Deep Convolutional Neural Networks." Remote Sensing of Environment 2024. [10]
Ma, Singh, and Zhao (2024) used dual-plyometric Sentinel 1 SAR data. This paper pointed out how using dual-polymetric data, especially in rough environments could actually help.

Limitation in Existing System

The current systems employed for oil spill detection have a number of problems. They include unavailability or poor quality of annotated data, which is essential for training deep learning models. Class imbalance among data, due to the vast difference between non–spill and spill images, resulting in a bias. Current systems also have problems generalizing over multiple environments, which could be because of low quality of data. In addition to that, satellite imagery is always prone to factors such as weather, cloud cover etc. These issues point out the need for a robust model and good quality data.

3. Methodology

In this current work, we highlight the significance of using advanced deep learning models to accurately detect oil spills using satellite images. Our work incorporates the utilization of U-Net, DeepLabV3, and PSPNet, all of which are renowned for their effectiveness in semantic segmentation tasks. U-Net, initially proposed for medical image segmentation, is highly effective in producing accurate pixel-level segmentations, a key aspect for oil spill detection. DeepLabV3, on the other hand, utilizes atrous convolutions to extract multiple scale contextual information and improve the performance of the model in detecting oil spills of diverse sizes. PSPNet utilizes pyramid pooling to pool global context and thus perform better in handling complex environmental changes in oil spill images. By using such advanced models we want to enhance the reliability and accuracy of detection of spills, finally leading to a safe system.

3.1 Dataset Overview:

The data used in this research is sourced publicly from a Kaggle dataset named - “oil-spill”. Figure 1 Oil spill dataset It contains high-resolution satellite images from different weather conditions and geographic locations, making it an ideal resource to train a deep-learning model over it. The dataset is also accompanied with ground-truth labels and pre-processed making it useful for segmentation processes. Variations in spill types, sizes and shapes could pose up as a challenge or a chance to train the model better.

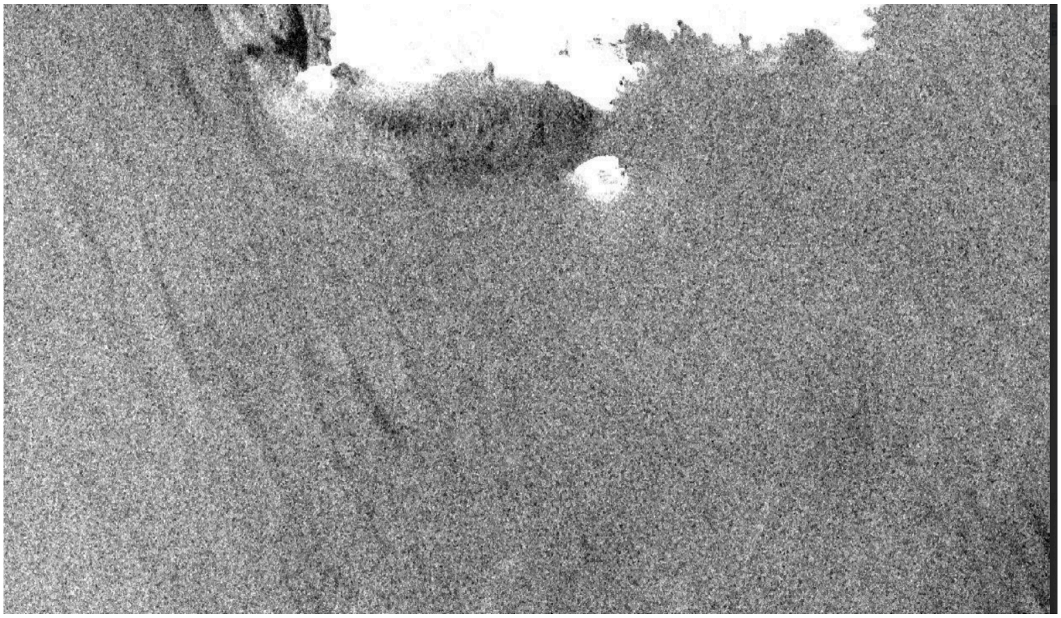

Figure 1: Oil Spill Dataset

3.2 Model Architectures:

Figure 2: Architecture Diagram compares the three deep learning models employed for oil spill detection: U-Net, DeepLabV3 and PSPNet. Each model has a unique feature that enables it to segment oil spills correctly. U-Net has an encoder-decoder architecture with skip connection, making it preserve spatial information. DeepLabV3 has a Resnet/Xception backbone combined with Atrous Spatial Pyramid Pooling (ASPP) making it better able to detect oil spills with different sizes. PSPNet, uses Pyramid Pooling Module (PPM) to enhance segmentation for images having different spill sizes. The below architecture Figure 2: Architecture Diagram displays the steps involved in the processes of each model, showing how the input image goes through all the models. Our goal is to compare these architectures and figure out which model outperforms the other and is the best to work in the case of oil spill detection.

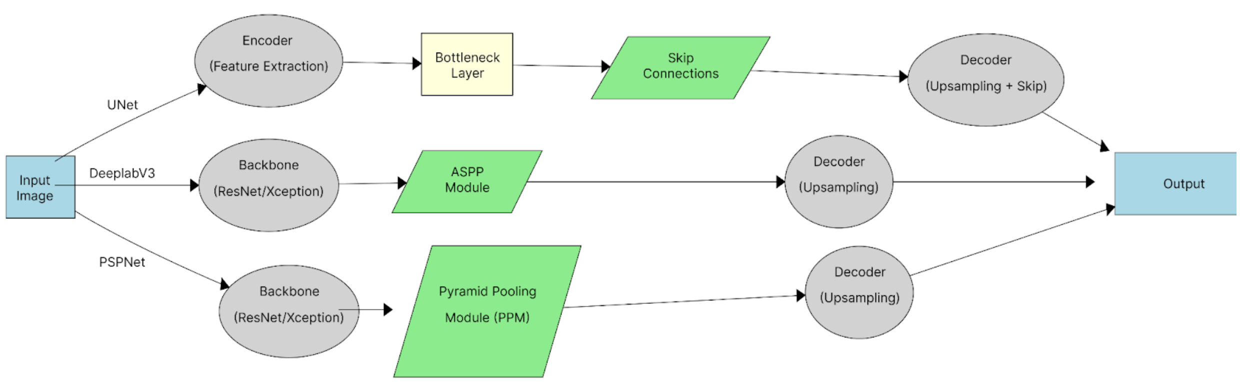

Figure 2: Architecture Diagram

3.3 U-Net Model

U-Net works best in the case of biomedical image segmentation like in Figure 3 where segmentation is done with a symmetric encoder-decoder structure and skip connections. The encoder path captures the context, and the decoder path achieves accurate localization. The structure is made up of two primary parts: the contracting path (encoder) and the expansive path (decoder). The contracting path is the standard configuration of a convolutional network in which two 3 x 3 convolutions (unpadded) are applied repeatedly, each of which is followed by an activation with a ReLU and a 2 x 2 max pooling operation with stride two for downsampling. Mathematically, this can be expressed as:

$$h_{enc}^{l+1} = MaxPool(\sigma(W_l * h_{enc}^{l} + b_l)) \quad (11)$$

where W_l and b_l weights and biases of the l^{th} layer, and σ denotes the ReLU activation function. In the expansive path, the features are sampled-up and combined with the high-resolution features from the contracting path using the concept of skip connections. This step can be described by:

$$h_{dec}^{l} = Concatenate(h_{dec}^{l+1}, h_{enc}^{l}) \ (12)$$

$$h_{dec}^{l-1} = \sigma(W_{up}^{l} * h_{dec}^{l} + b_l) \ (13)$$

where W_{enc}^{l} is the up-convolution (transposed convolution) weights. The architecture enables U-Net to utilize both local context from the contracting path and contextual information from the expansive path, making it very efficient for segmentation tasks with the need for accurate boundary delineation.

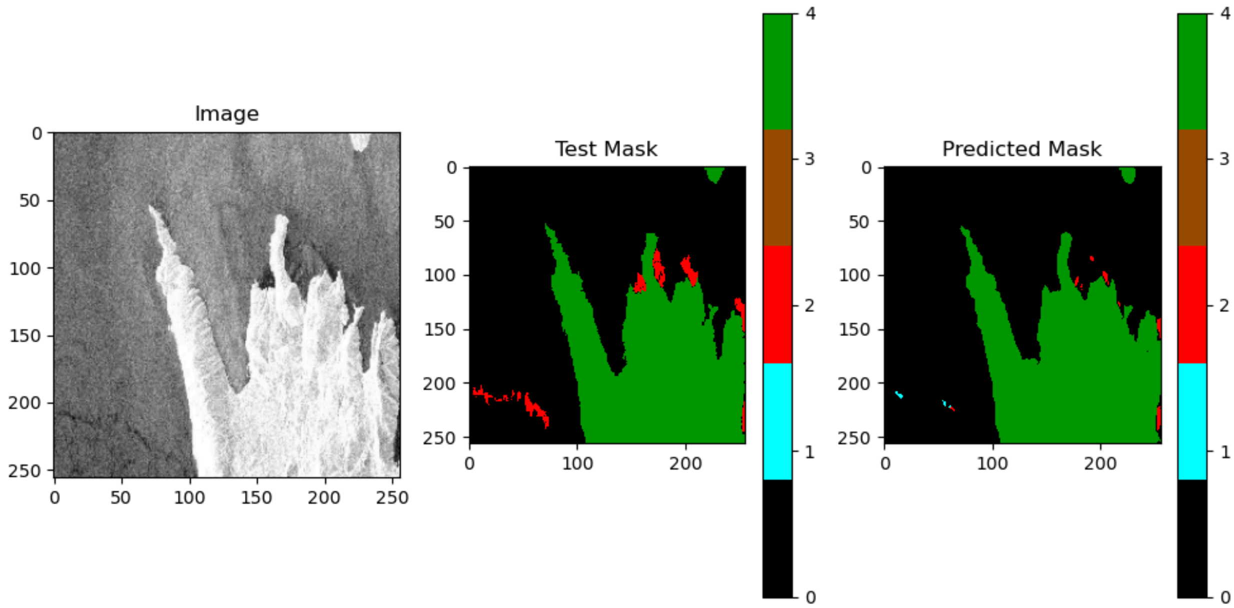

Figure 3: Segmentation Results Using U-Net

3.4 DeepLabV3

DeepLabv3 uses atrous (or dilated) convolutions to modulate the resolution at which features are calculated. Atrous convolution adds holes (or zeros) in between filter elements so that there is a larger field of view without an increase in parameters, see Figure 4 to see predicted visualizations. The operation of atrous convolution can be represented using the following equation:

$$y[i] = \sum x[i + r.k]w[k] \ (2)$$

where r is the atrous rate, x is the input signal, y is the output signal, and w is the filter. This approach efficiently expands the view field of filters without adding parameters or computational expense.

In addition to atrous convolutions, DeepLabv3 utilizes Atrous Spatial Pyramid Pooling (ASPP) to capture multi-scaale context information by applying multiple atrous convolutions with different rates. ASPP applies paralllel dilated convolution with varying rate to feel an incoming convolutional feature layer with filters at multiple sampling rates, capturing objects and image context at different scales. The ASPP operation can be expressed as:

$$ASSP(x) = Concatenate[conv\,1 \times 1(x), conv3 \times 3_{r}(x), conv\,3 \times 3_{2r}(x), conv3 \times 3_{3r}(x)]$$

where r denotes the atrous rate.

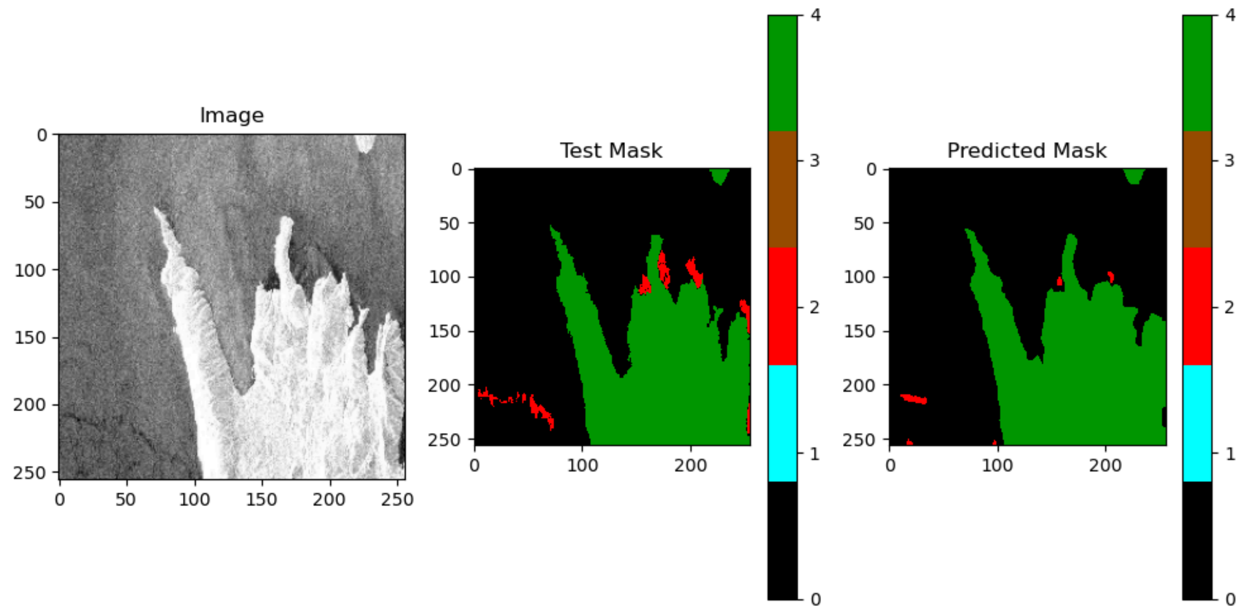

Figure 4: Segmentation Results Using DeepLabV3

3.5 PSPNet

PSPNet (Pyramid Scene Parsing Network) as you see in Figure 5 uses a pyramid pooling module to extract multi-scale context through pooling features in various scales. This enables the network to summarize contextual information with varying levels of granularity. The pyramid pooling module splits the image into regions with varying sizes and extracts both global and local context. The various levels of pooling operations can be formulated as:

$$P_k(x) = Pooling(x, scale_k) \quad (4)$$

where $P_k(x)$ refers to the pooled feature at scale k and $scale_k$ represents the corresponding spatial resolution at the *k-th* pooling level. Following pooling at various levels, the features are concatenated to generate a rich, multi-scale feature representation, which increases the model's capability to perceive complex scene structures.

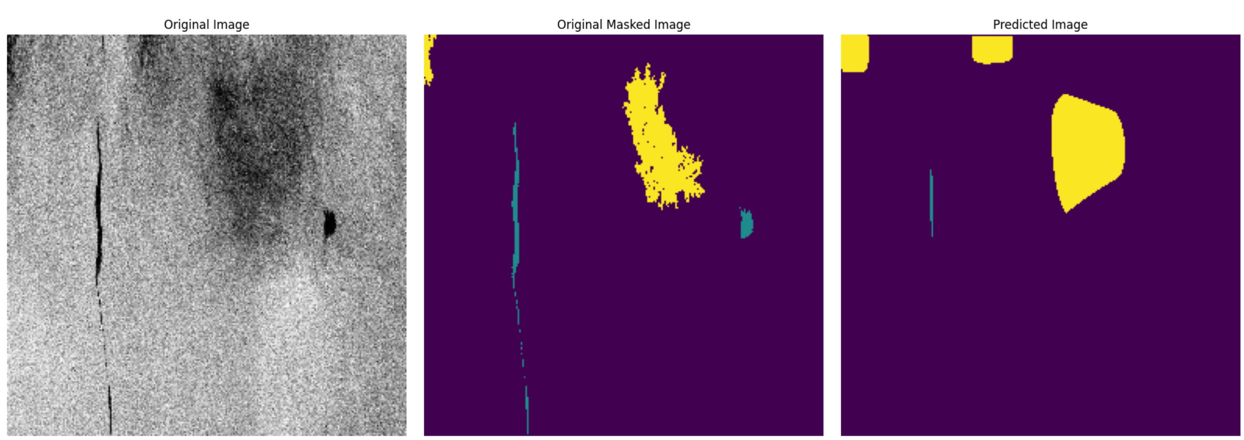

Figure 5: Segmentation Results using PSPNet

Research Gaps:

Oil spill detection research has several gaps, such as limited model generalizability over different environmental settings, lack of large, annotated datasets, difficulty in identifying small spills or those against complex backgrounds, and insufficient models that are resilient to noise and artifacts in satellite imagery, along with lacking real-time scalability for large areas of monitoring. Multi-modal and multi-spectral data integration, enhanced model explainability and interpretability, and developing higher-level evaluation metrics than simple accuracy are areas for future investigation.

4. Results:

4.1. Performance Metrics

DL Model	Precision	Recall	F1 Score
UNet	0.9462	0.9499	0.9475
DeepLab V3	0.9534	0.9563	0.9543
PSPNet	0.8932	0.8997	0.8993

Table 1: Model Performance Metrics

As shown in Table 1, the efficiency of the three models, U-Net, DeepLabV3, and PSPNet, in oil spill detection was compared with precision, recall, and F1 score being the primary metrics for comparison. These metrics lead us to comparing the models and their capability to successfully detect spill areas, reducing false positives and negatives.

U-Net showed a precision of 94.62%, a recall of 94.99%, and an F1 score of 94.75%. These findings show that U-Net was good at perfectly classifying most spill areas with a minimum amount of false positives. The recall score also confirms this on how it is a decent model of oil spill detection. The main advantage of it is that it takes the least time compared to all the other models.

DeepLabV3 showed better results compared to U-Net, a precision of 95.43%, a recall of 95.63%, and an F1 score of 95.43%. DeepLabV3's increased precision and recall reflect its capacity to both accurately detect oil spills (precision) and detect most of the true oil spill regions (recall) better than U-Net. This improved performance is due to its deeper and more advanced architecture, which includes dilated convolutions and a broader context for improved pixel-wise segmentation, particularly for small and intricate spill shapes.

PSPNet is presumed to have done the worst. Considering PSPNet's Pyramid Scene Parsing mechanism, which is meant to extract multi-scale contextual information, it has given a nice trade-off between precision and recall. PSPNet had a precision of 89.32%, a recall of 89.97%, and an F1 score of 89.93%. Though a bit behind UNet, PSPNet's multi-scale strategy did a good job, especially for big oil spills.

4.2. Training & Validation Performance

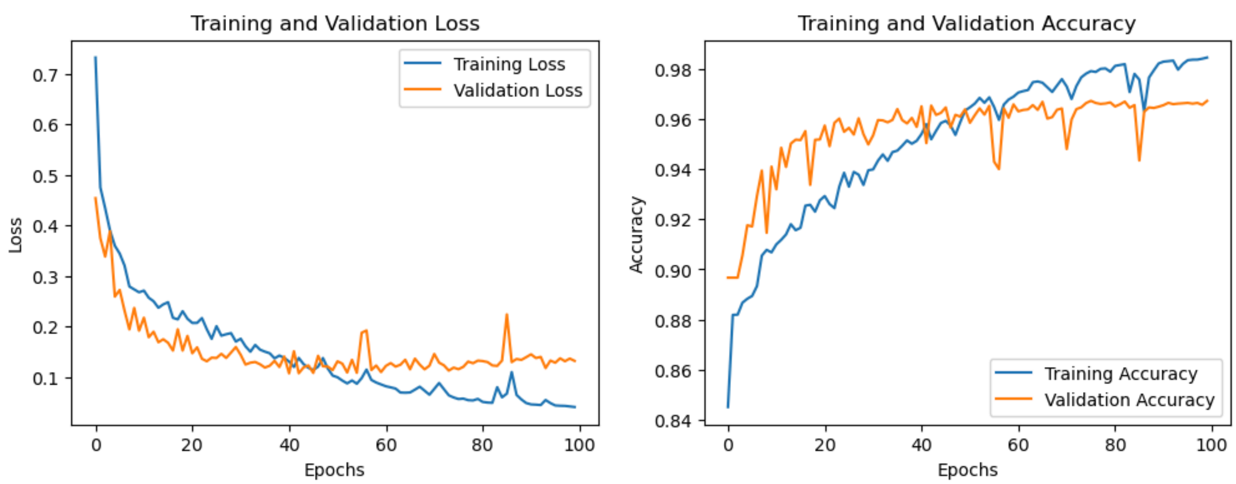

Figure 6: U-Net Training & Validation Performance

Training and validation loss graphs for the two DeepLabV3 and U-Net demonstrated a consistent decline in loss with each epoch, proving successful learning (refer to Figure 6 for U-Net and Figure 7 for DeepLabV3). There was evidence of overfitting on both models, especially during later training stages when the validation loss increased minimally while the training loss kept on improving. This implies that the models may have over-specialized to the training data, and other regularization strategies (like dropout or data augmentation) could help improve generalization to unseen data. The overfitting was more extreme in U-Net.

The training and validation accuracy curves also reflected robust performance. DeepLabV3, however, reflected a marginally higher accuracy than U-Net, which is consistent with its better precision and recall values (see Figure 6 for U-Net and Figure 7 for DeepLabV3). DeepLabV3 attained training accuracy of 98.5% and validation accuracy of 97.3%, whereas U-Net attained training accuracy of 97.2% and validation accuracy of 96.1%.

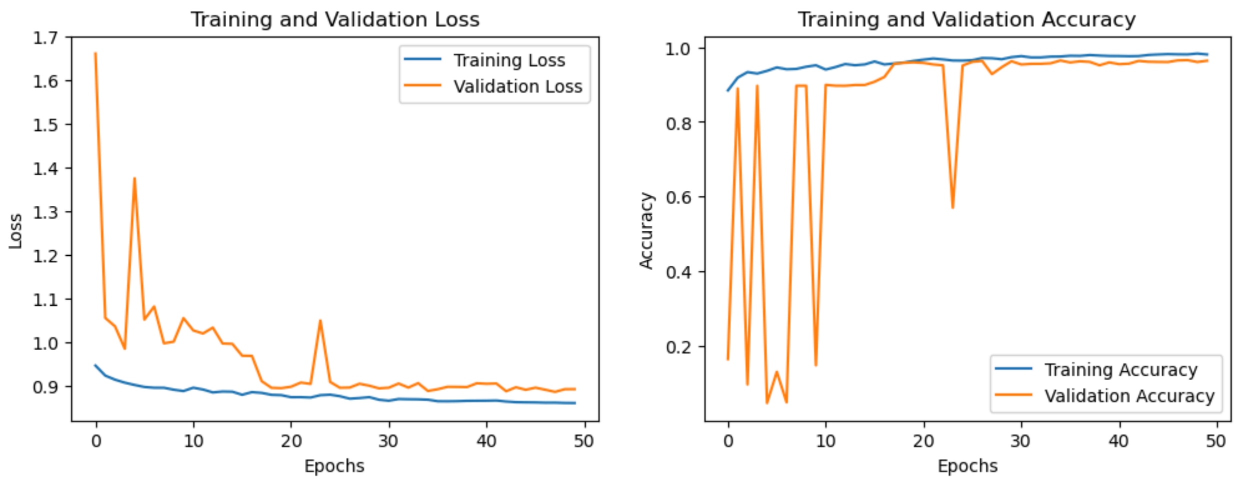

Figure 7: DeepLab V3 Training & Validation Performance

Figure 8 shows the plot indicates the training and validation loss of PSPNet for 50 epochs. The loss drops sharply in the initial 10 epochs and becomes stable at epoch 15, reflecting good learning. Minor fluctuations in validation loss after epoch 30 reflect minor overfitting. In spite of this, the model has a high F1-score of approximately 89%, reflecting good segmentation performance.

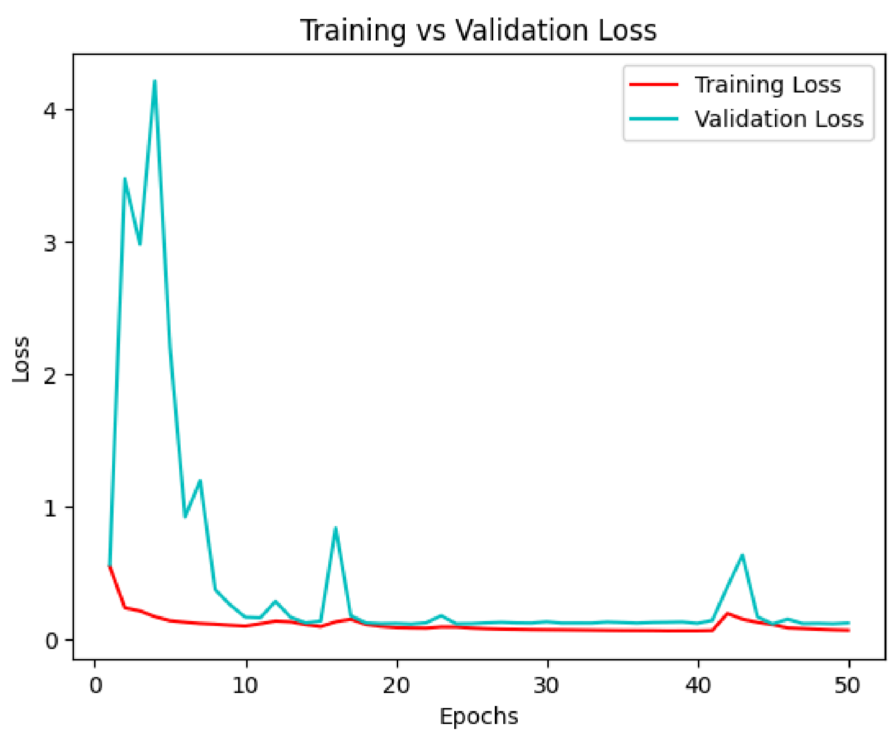

Figure 8: PSPNet Training & Validation Performance

4.3 Qualitative Results:

DeepLabV3 always produced sharper and more accurate boundaries of oil spill areas with fewer false positives than U-Net. U-Net, although still precise, occasionally identified small patches of water as oil spill areas, leading to slight false positives.

On the other hand, PSPNet, although coming second to UNet, also showed great segmentation performance, especially for bigger oil spills. Its ability to understand contexts at multiple scales helped it to deal with spills of different sizes well, but it could not perform as accurately in detecting smaller oil spills.

Conclusion:

The comparison between U-Net, DeepLabV3, and PSPNet emphasizes that even though U-Net is an effective and trustworthy model for oil spill detection, DeepLabV3 yields better results in terms of our comparison metrics, particularly regarding complex and small oil spill detection. The ability of DeepLabV3 to pay attention to multi-scale context and precise pixel-level predictions resulted in its enhanced performance. Nevertheless, all models, including PSPNet, continued to struggle with noisy data and intricate backgrounds, where false positives and false negatives were sometimes present.

In short, DeepLabV3 performs better than U-Net and PSPNet for oil spill detection according to the metrics considered. Nevertheless, U-Net is a robust model with high recall and can be used in situations where every possible oil spill must be captured. PSPNet, being multi-scale, provides competitive performance, especially for large spill detection. Additional research can concentrate on fine-tuning these models to perform well for small spills, noisy data, and robustness under different environmental conditions.

Future Work:

Future research can investigate combining attention mechanisms like CBAM or SE blocks for improved feature extraction and segmentation accuracy, and can also integrate AIS (Automatic Identification System) information to aid in the detection of oil spills by correlating the detected spills with adjacent ship movements.

References:

1. Solberg, A., Brekke, C., and Husoy, P. "Oil Spill Detection in Radarsat and Envisat SAR Images." IEEE Transactions on Geoscience and Remote Sensing 2007.

2. Krestenitis, M., Oikonomou, M., and Kalogiros, J. "A Benchmark Dataset for Oil Spill Detection Using Deep Learning." Remote Sensing 2019.

3. Abdimanap, G., Omarov, Y., and Abdallah, A. "ROSID: A Remote Sensing Dataset for Onshore Oil Spill Detection." Environmental Monitoring and Assessment 2023.

4. Mahmoud, H., Zhang, J., and Kim, T. "A Dual Attention Model for SAR-Based Oil Spill Detection Using Deep Learning." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2022.

5. Li, Z., Wang, S., and Liu, Q. "Object-Oriented Oil Spill Detection from SAR Images Using Adaptive Local Thresholding and Fuzzy Classification." International Journal of Remote Sensing 2018.

6. Guo, H., Xu, Y., and Chen, L. "CNN-Based Oil Spill Identification Using Polarimetric SAR Data." IEEE Transactions on Geoscience and Remote Sensing 2020.

7. Shaban, M., Al-Ali, M., and Habib, A. "A Two-Stage Deep Learning Approach for Oil Spill Detection in Highly Unbalanced SAR Data." ISPRS Journal of Photogrammetry and Remote Sensing 2021.

8. Fan, Y., Zhao, P., and Sun, W. "Feature Merge Network for Oil Spill Detection Using Synthetic Aperture Radar Imagery." Journal of Environmental Management 2022.

9. Rousso, R., Patel, K., and Meena, S. "Enhancing Oil Spill Detection with Image Filtering and Deep Learning Models." Geosciences 2023.

10. Ma, L., Singh, V., and Zhao, T. "Dual-Polarimetric Sentinel-1 SAR for Oil Spill Detection Using Deep Convolutional Neural Networks." Remote Sensing of Environment 2024.