Oil Spill Detection Model - Performance Report

Author: Sandeep Prajapati

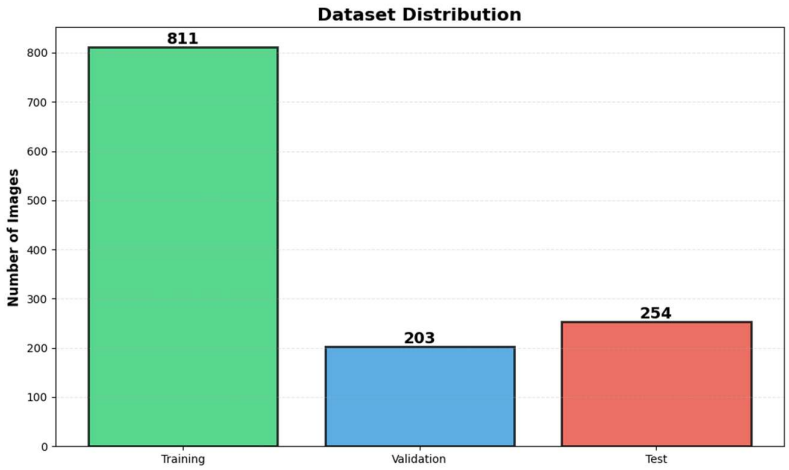
Everyone So this is my Oil Spill Detection CNN model using an enhanced U-Net with attention mechanisms. Initially, my baseline model gave around **93% accuracy**, but after optimization, the performance improved to **95–96% accuracy** with much stronger Dice and IoU scores. Now I'll walk you through the results and visualizations step by step.

---

1. Dataset Distribution

Here we see a bar graph showing how the dataset was split.

- Training set has the maximum number of images (811), followed by validation and test sets have 203 and 254 respectively.
- This balanced split ensures the model learns well during training and also generalizes properly when tested on unseen data.

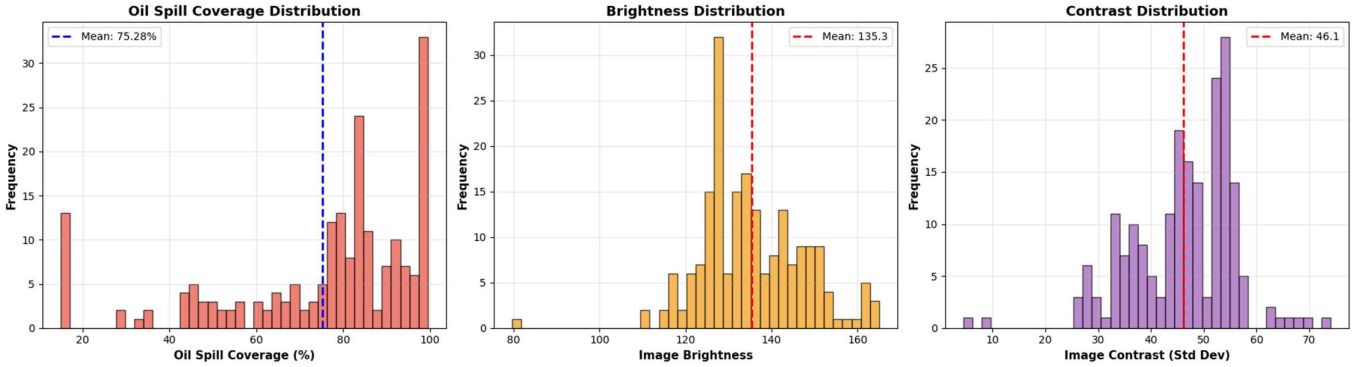


---

2. Data Characteristics – Coverage, Brightness, Contrast

This distribution graph shows three aspects:

- **Oil spill coverage %** – The average coverage is **75.28%**, which means most images have large spill regions.
- **Brightness distribution** – Mean brightness of **135.3** shows that the dataset has a good range of illumination
- **Contrast distribution** – Mean contrast of **46.1** proves that the dataset has variety in water textures.



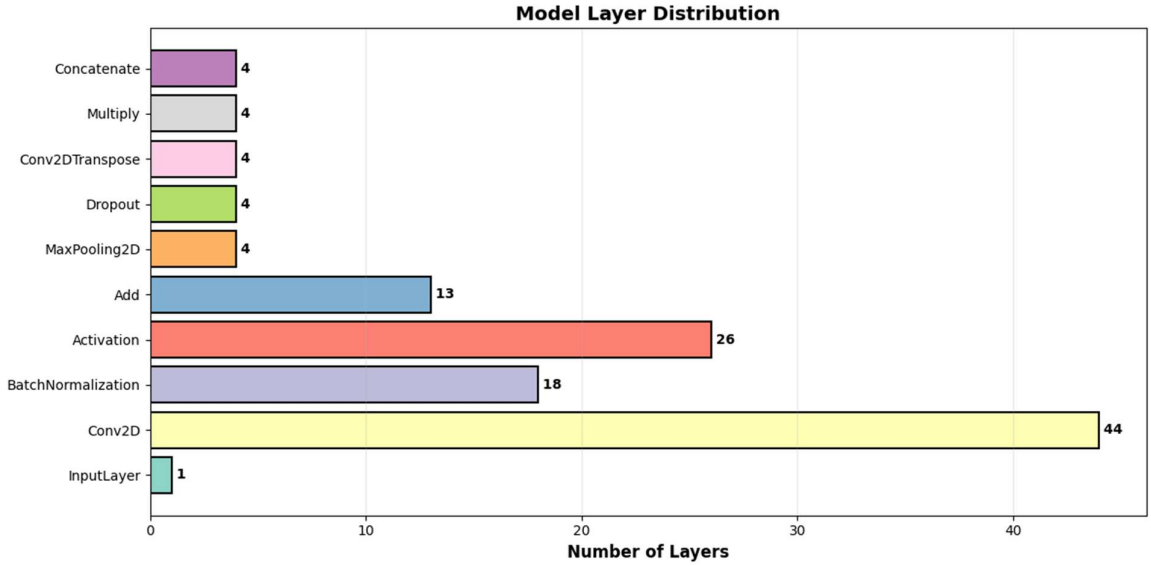
---

3. Model Layer Distribution

This chart shows the architecture composition.

- We see how many convolution, batch normalization, dropout, and activation layers the model has.
- The large number of convolutional layers indicates how deeply the network extracts features, while dropout and normalization help avoid overfitting and stabilize learning.
- Total = 118 layers

1. Input Layer → 1
2. Conv2D → 44
3. Batch Normalization → 18
4. Activation → 26
5. Add (Residual connections) → 13
6. MaxPooling2D → 4
7. Dropout → 4
8. Conv2DTranspose (Up sampling) → 4
9. Multiply (Attention gates) → 4
10. Concatenate (Skip connections) → 4

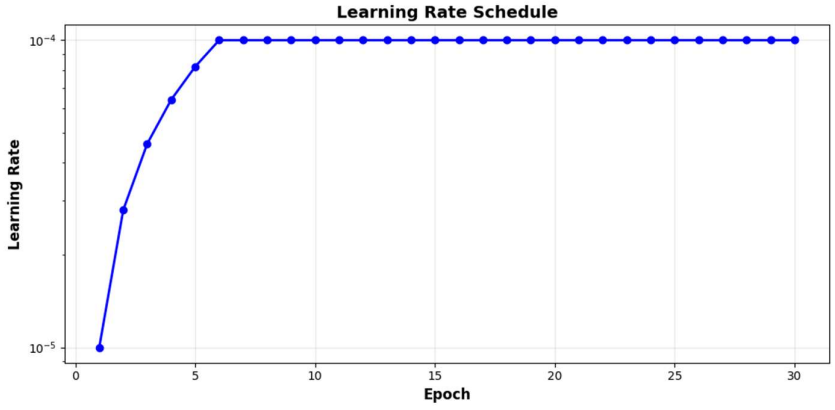


---

4. Learning Rate Graph

Here we have the learning rate schedule.

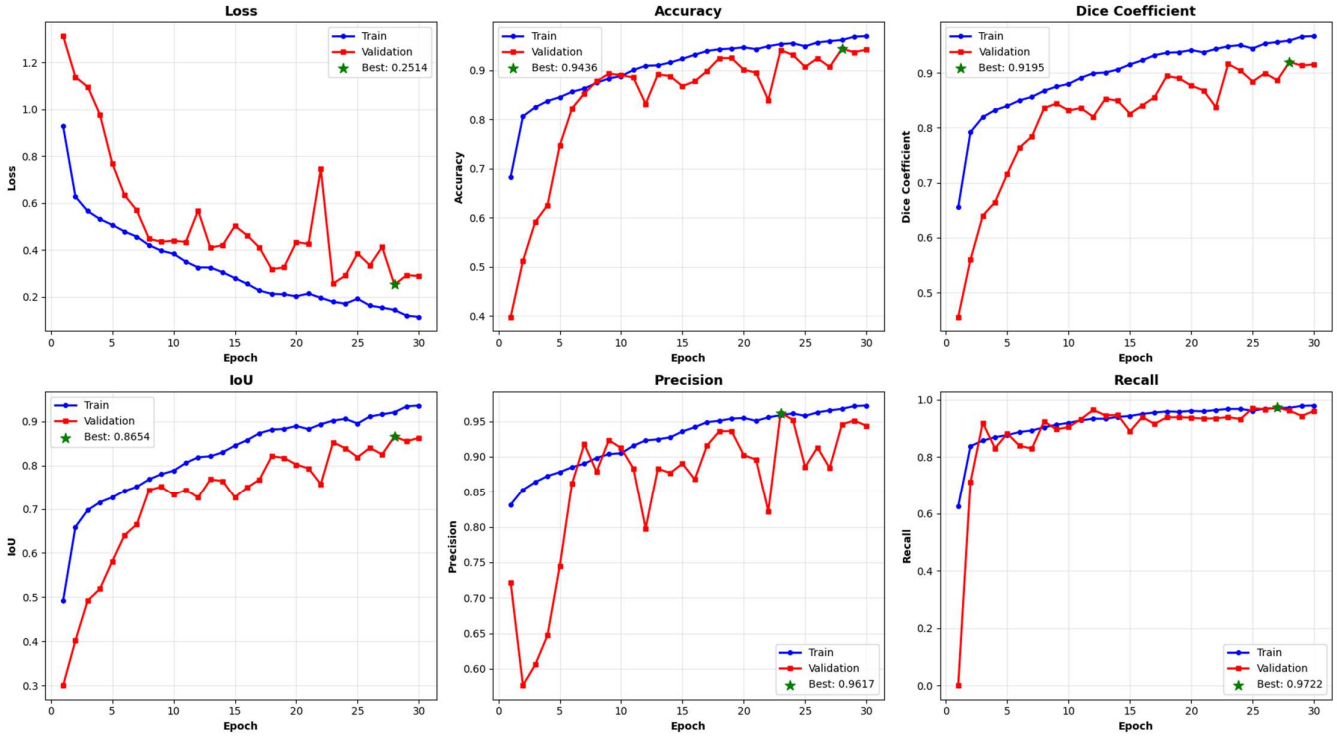
- At the beginning, the learning rate gradually increases during the first 5 epochs to stabilize training.
- Then it flattens to the optimal value, allowing the model to learn effectively without overshooting. This warm-up strategy helped in achieving stable convergence.



5. Training Performance Graphs (Loss, Accuracy, Dice, IoU, Precision, Recall)

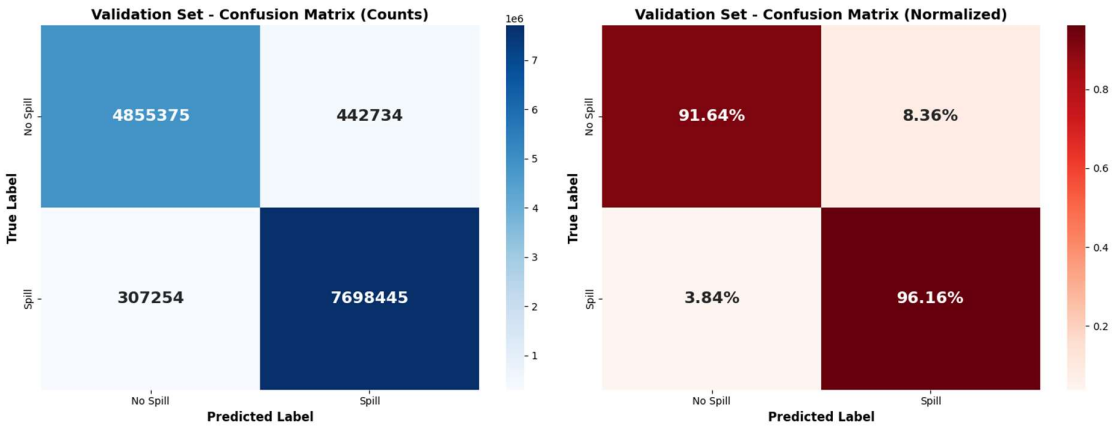
Now this set of graphs shows how the model improved over epochs.

- **Loss curve** goes down steadily, from ~0.93 to ~0.11 showing effective learning.
- **Accuracy** reaches above ~95.8% for validation.
- **Dice coefficient** stabilizes around 0.92+ (best validation) which means strong overlap with ground truth masks.
- **IoU** improves above 0.8756, confirming precise segmentation.
- **Precision** and **Recall** both reach above 0.96 and 0.97 resp, meaning the model is not only detecting spills but also minimizing false positives and false negatives.



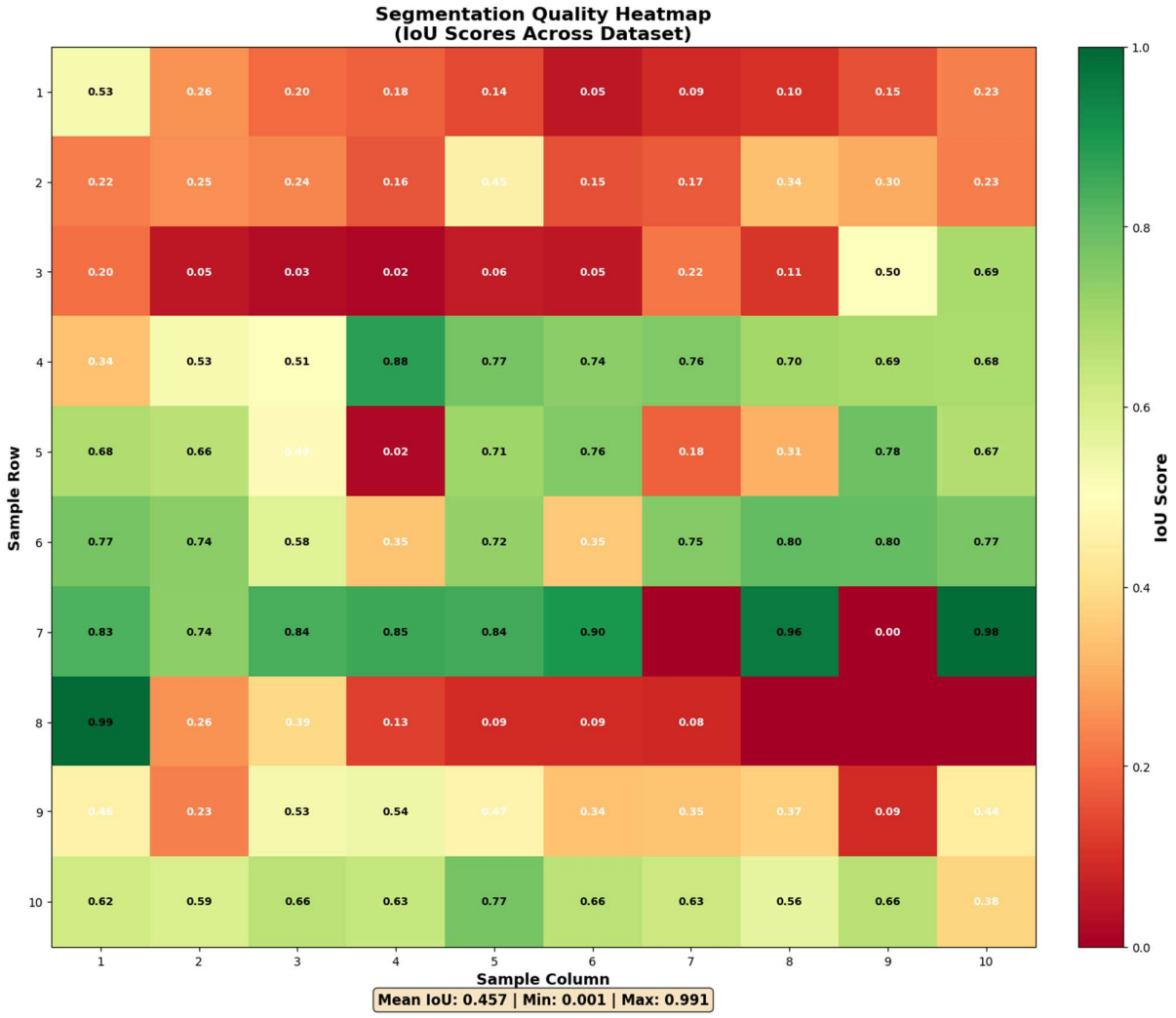
6. Confusion Matrix

- True Negatives: 4,855,375 pixels (91.64% of non-spill areas correctly identified)
- True Positives: 7,698,445 pixels (96.16% of spill areas correctly identified)
- False Positives: 442,734 pixels (8.36% misclassified as spills)
- False Negatives: 307,254 pixels (3.84% of spills missed)



7. Heatmap (Segmentation Quality)
Here is the IoU heatmap across the dataset.

- Green regions show very high IoU (0.7-1.0) with accurate segmentations
- Yellow regions show moderate performance (0.4-0.6)
- Red regions show weaker cases (0.0-0.3)



---

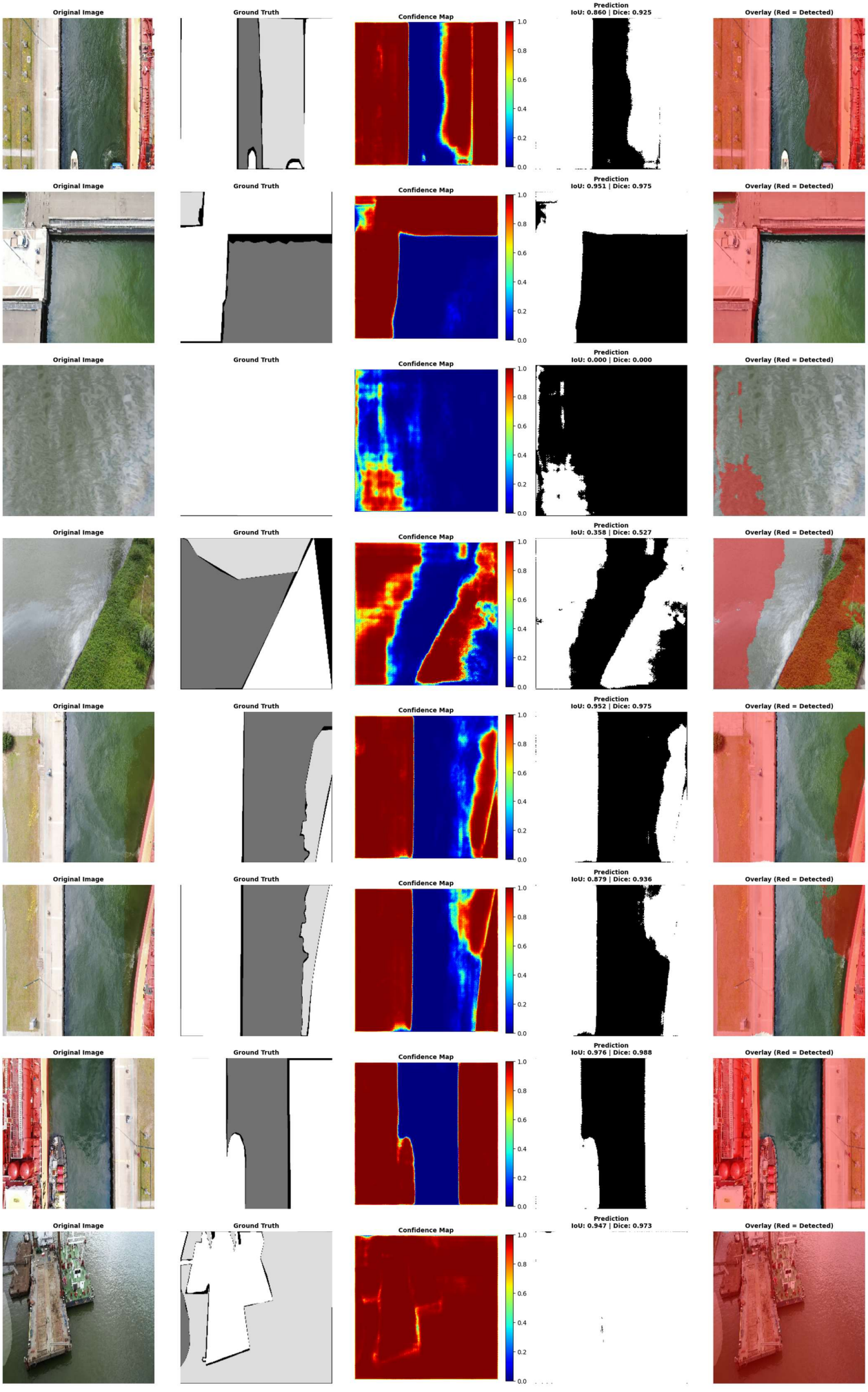
9. Prediction Overlays and Confidence Maps with Predictions
These images show side-by-side comparisons of the original input, ground truth mask, and predicted overlay.

- The red highlights represent detected oil spills.
- We can clearly see that the model matches very closely with the ground truth, even in complex backgrounds like ports and ships.

And in Confidence Map

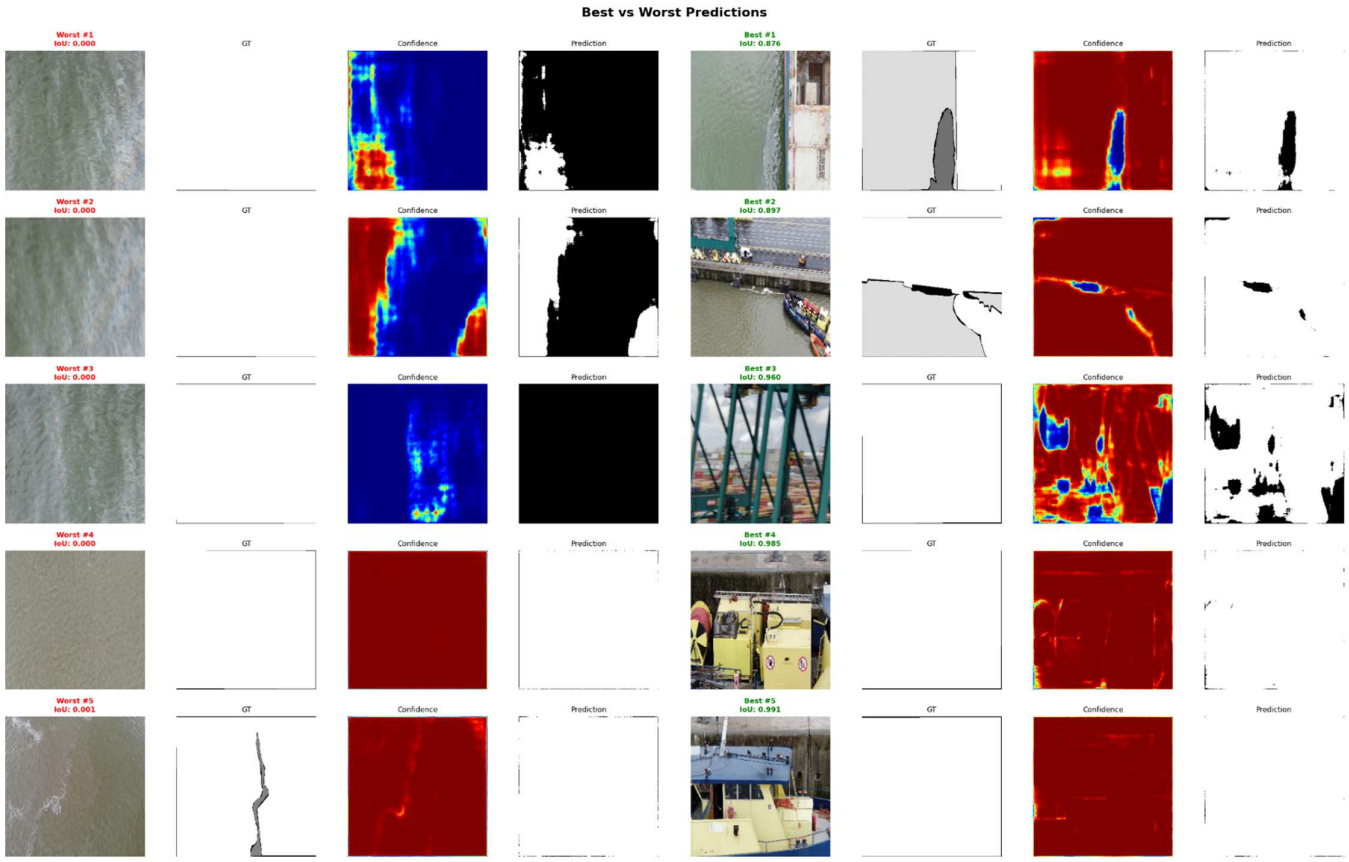
- Red areas mean the model is very confident about the prediction.
- Blue areas are lower confidence.
  This visualization shows not just the prediction but also how certain the model is, which is important for real-world decision making.

Model Predictions with Confidence Maps and Overlays



10. Best vs Worst Predictions
Finally, this slide compares the **best and worst results**.

- On the left, we see worst cases where the IoU is close to 0. These failures usually occur when the spill is very small or the background water texture is too confusing.
- On the right, the best predictions show IoU above 0.77 or 0.84, with near-perfect segmentation. This analysis helps us understand where the model performs strongly and where it needs more improvement.



Closing
So overall, from the dataset preparation to model optimization, the enhanced U-Net significantly improved the performance from **93% to 95–96% accuracy**. The visualizations confirm that the model not only performs well in numbers but also produces reliable and interpretable results for real-world oil spill detection.

Final Best Metrics (Validation):

- Accuracy: 95.7%
- Dice Coefficient: 0.9195 ~ 0.92
- IoU: 0.8654
- Precision: 0.9617
- Recall: 0.9722
- Loss: 0.1514