

Alberta Wells Dataset: Pinpointing Oil and Gas Wells from Satellite Imagery



Pratinav Seth  **Michelle Lin**  **Brefo Dwamena Yaw**  **Jade Boutot**  **Mary Kang**  **David Rolnick** 

★ - Denotes co-first authorship

PROBLEM: Millions of abandoned oil and gas wells emit methane and contaminate groundwater, and many well locations are unknown, preventing remediation efforts.

SOLUTION: Leveraging remote sensing and machine learning to detect wells from medium-resolution satellite imagery.

DATASET CREATION:

- Data preprocessing & cleaning followed by domain expert-verified filtering and quality control.
- Create non-overlapping image patches (1.1025 km² per patch) and corresponding labels based on domain knowledge.
- Dataset splitting: A two-stage K-means clustering approach to group well locations into regional clusters before splitting into train, validation, and test sets. Ensures diverse geographic representation in each split while minimizing data leakage.

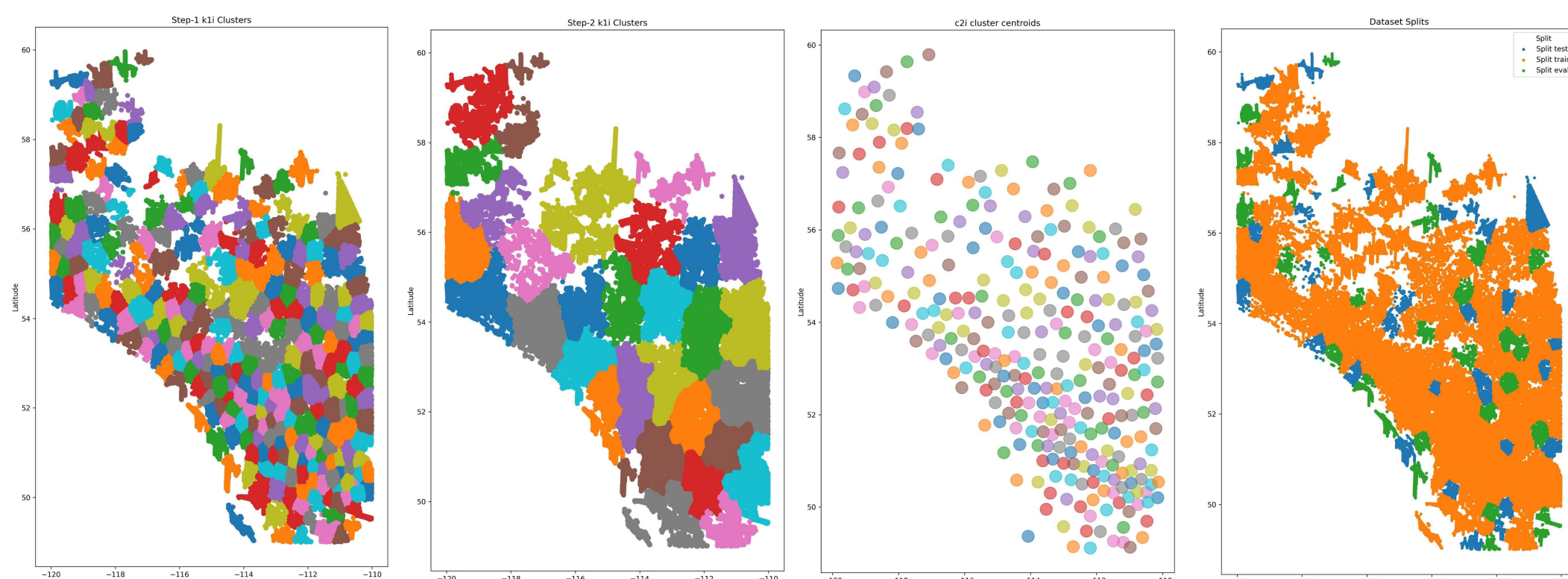


Figure 2: Illustration of the outcome of applying our dataset splitting algorithm

EXPERIMENTS & ANALYSIS:

- Evaluated models for detection and segmentation task.
- Binary segmentation (Best: U-Net with EfficientNet-B6)
- Object detection (Best: DETR with ResNet50)
- Small object size (30 pixels) and large number of wells in a single instance makes task challenging.
- Training on only active wells misses abandoned ones; including active, suspended, and abandoned wells improves generalization.

Architecture	Backbone	IoU	F1 Score	Precision	Recall
U-Net	ResNet50	58	61.9	90.2	62.3
U-Net	EfficientNetB6	60.4	64.8	87.8	66.3
DeepLabV3+	ResNet50	56.8	60.6	89.4	61.8
Segformer	mit-b0-ade	57.6	61.3	82.6	69.2
UperNet	ConvNext-Base	59.7	63.8	81.1	72.2
UperNet	swin small	59.9	64.2	80.6	73.1

Table 1: Results for the binary segmentation task for a variety of models evaluated over the test set. We report the Intersection over Union (IoU), F1 Score, Precision and Recall.

Architecture	Backbone	IoU _{0.1}	IoU _{0.3}	IoU _{0.5}	mAP ₅₀	mAP _{50:95}
FasterRCNN	ResNet50	36.79	46.95	61.29	5.2	19.12
FCOS	ResNet50	34.79	48.51	62.66	9.67	30.46
SSD Lite	MobileNetV2	33.91	50.3	65.07	9.76	25.14
DETR	ResNet50	41.78	51.15	63.17	15.22	38.45

Table 2: Results for the object detection task for a variety of models evaluated over the test set. We report the intersection over union (IoU) over thresholds (0.1,0.3,0.5) and the mean average precision (mAP) for both IoU = 0.5 and IoU ∈ [0.5,0.95] thresholds.

ALBERTA WELLS DATASET (AWD):

- First large-scale remote sensing benchmark dataset (213,447 wells) for oil & gas well detection.
- High resolution (3m/px) imagery from Planet Labs (RGB + NIR). with diverse landscape (prairies, forests, mountains).
- Diverse state of operation (active, abandoned, suspended) of wells in the dataset.

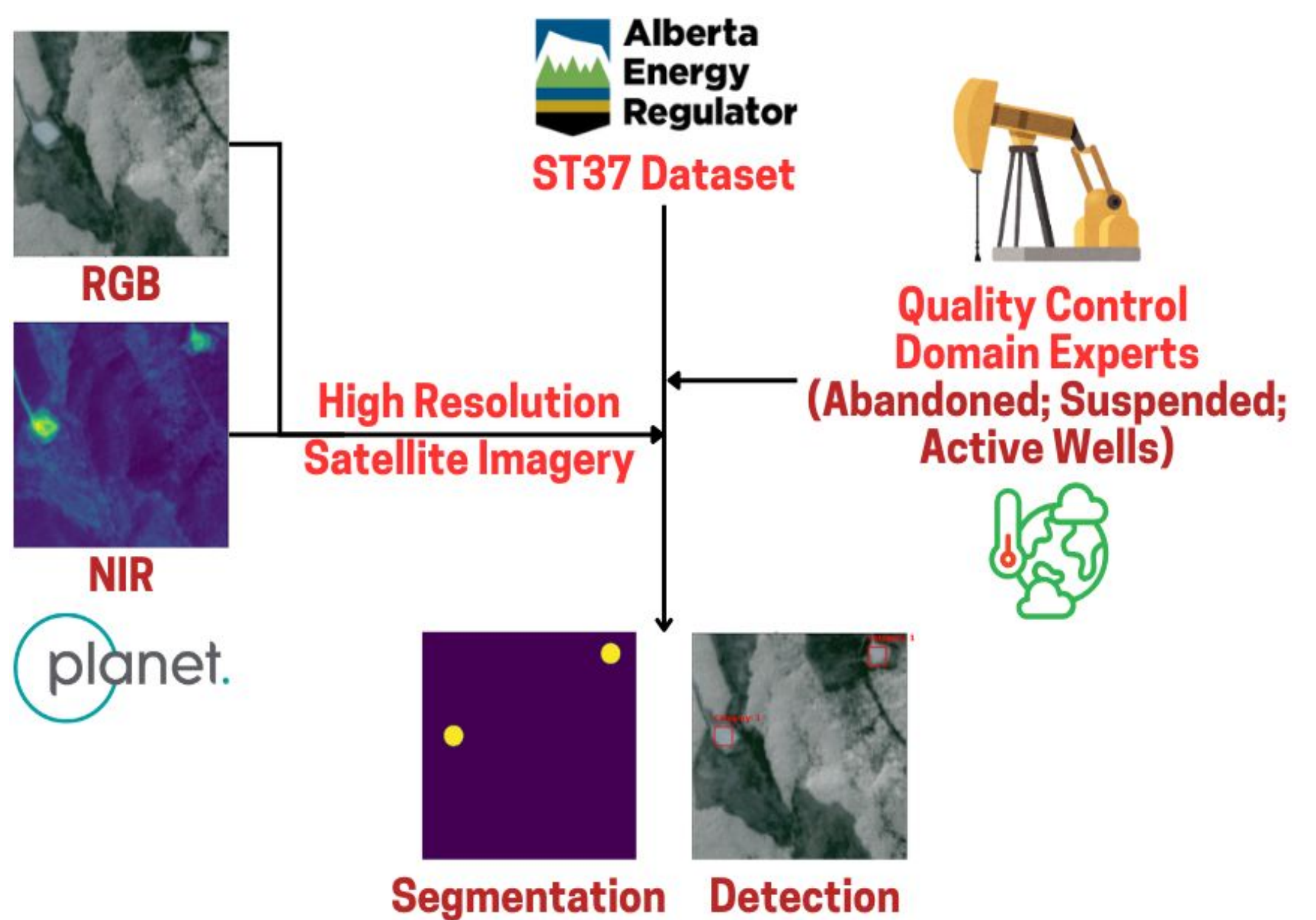


Figure 1: Alberta Well Dataset Overview

QUALITATIVE SAMPLES:

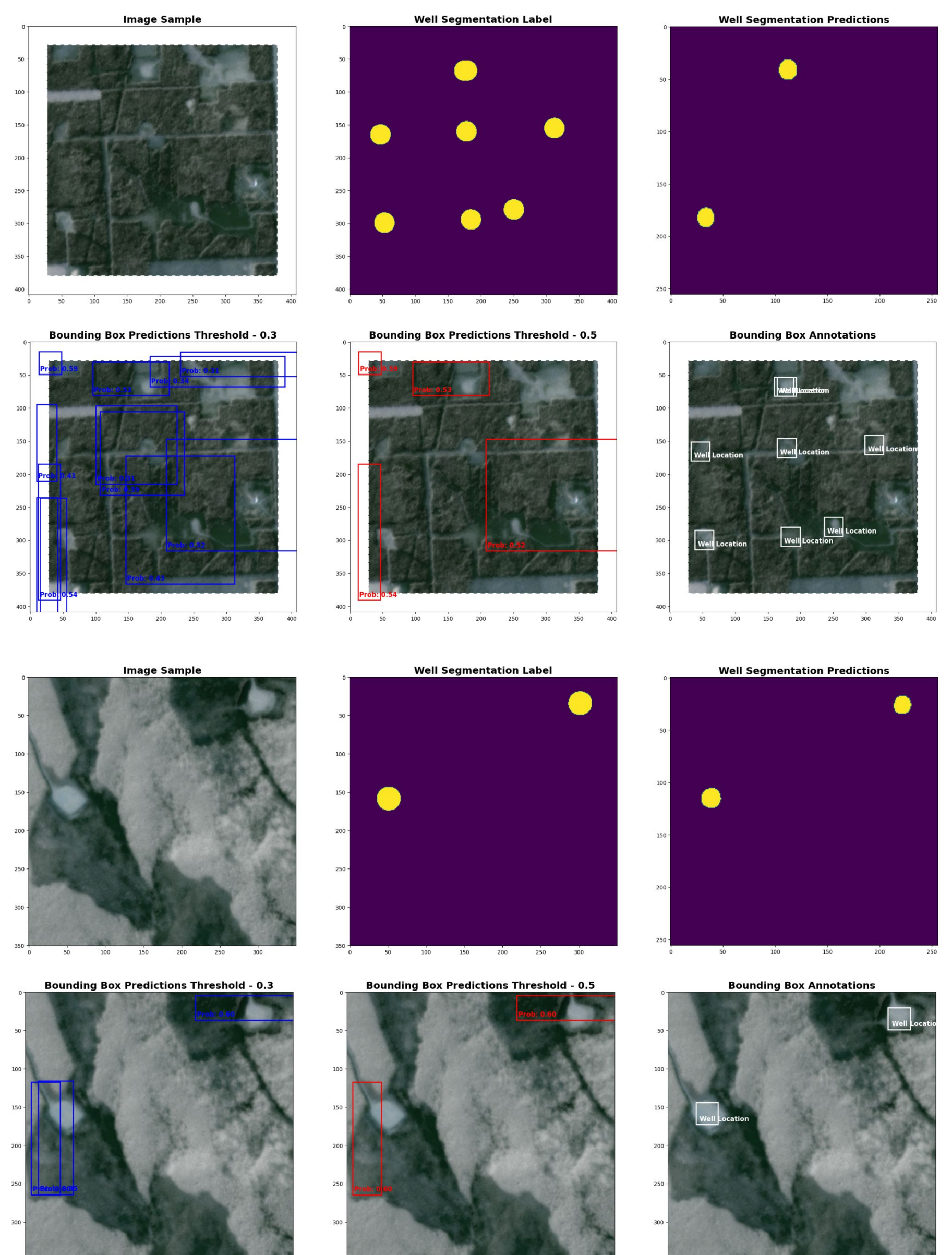


Figure 3: Sample image patches from our dataset illustrates few and multiple wells and corresponding qualitative results with predictions.