SmokeU-Net: Smoke Column Segmentation from Satellite Imagery using Deep U-Net Architecture

Arya Shah
Asian Institute of Technology
Department of Computer Science and Information Management
st125462@ait.asia

Abstract

This research proposes an automated deep learning approach for segmenting smoke columns in satellite imagery using U-Net architecture. Current manual and semi-automated methods for smoke plume detection are time-consuming and prone to human error. We aim to develop a robust segmentation model trained on multi-spectral satellite data that can accurately identify and segment smoke columns from wildfires. The proposed framework will leverage U-Net's encoder-decoder architecture to process multiple spectral bands and generate precise smoke column masks. Through this work, we expect to demonstrate significant improvements in smoke plume detection accuracy and processing efficiency compared to traditional threshold-based methods. This research will contribute to advancing automated wildfire monitoring capabilities and support early warning systems for disaster management.

1. Introduction

Wildfires have become increasingly devastating global phenomena, with their frequency and intensity escalating due to climate change. Satellite imagery plays a crucial role in wildfire monitoring and management, providing continuous coverage of vast geographical areas. Current operational systems rely heavily on manual interpretation or basic thresholding techniques to detect and track smoke plumes from these fires

Traditional methods for smoke plume detection in satellite imagery face significant challenges. Manual interpretation is time-consuming and subject to human fatigue and error, while simple threshold-based algorithms struggle with varying atmospheric conditions and often produce false positives. These limitations severely impact the speed and reliability of wildfire monitoring systems

Therefore: We propose a deep learning-based approach using U-Net architecture to automatically segment smoke columns in multi-spectral satellite imagery. Our framework will process multiple spectral bands to generate precise smoke column masks, leveraging U-Net's ability to capture both local and global contextual information through its encoder-decoder architecture.

The key findings we aim to demonstrate:

1. Development of a robust U-Net-based model for automated smoke column segmentation
2. Effective utilization of multi-spectral satellite data for improved detection accuracy
3. Comprehensive evaluation of the model's performance across varying atmospheric conditions
4. Comparative analysis against traditional threshold-based methods

The contributions of this work will be:

1. A novel deep learning framework specifically designed for smoke column segmentation in satellite imagery
2. A systematic evaluation methodology for assessing smoke detection accuracy in multi-spectral satellite data
3. An open-source implementation to support future research in satellite-based wildfire monitoring.

1.1. Research Questions

Our research addresses several fundamental questions in the domain of smoke column segmentation using satellite imagery:

1.1.1 Primary Research Question

- How effectively can a U-Net architecture perform binary segmentation of smoke columns in GOES-16 satellite imagery for wildfire monitoring?

1.1.2 Technical Research Questions

- What is the optimal configuration of U-Net hyperparameters (learning rate, batch size, number of layers) for achieving accurate smoke column segmentation while maintaining computational efficiency?
- How does the encoder-decoder architecture of U-Net capture both local features (smoke texture) and global context (overall smoke pattern) in single-channel satellite imagery?
- How does the choice of loss function and optimization strategy affect the model's ability to learn smoke column boundaries?

1.1.3 *Performance Research Questions*

- What level of segmentation accuracy can be achieved using a basic U-Net architecture with single-channel input compared to traditional thresholding methods?
- How does the model's performance vary with different types of smoke patterns (thin vs. dense, small vs. large plumes)?
- What is the trade-off between model complexity and segmentation accuracy in the context of smoke column detection?

1.2. Hypotheses

1.2.1 *H1: Architecture Performance*

The U-Net architecture, with its symmetric encoder-decoder structure and skip connections, will achieve segmentation accuracy exceeding 80% on GOES-16 satellite imagery due to its ability to:

- Preserve fine spatial details through skip connections
- Capture hierarchical features through multiple convolutional layers

Maintain spatial information through the decoder path

1.2.2 *H2: Feature Learning*

The encoder path of U-Net will effectively learn hierarchical features of smoke columns, where:

- Early layers will capture basic edge and texture information
- Middle layers will learn smoke pattern structures
- Deep layers will understand global context and smoke distribution

This hierarchical learning will result in more accurate boundary detection compared to traditional methods.

1.2.3 *H3: Training Convergence*

The combination of binary cross-entropy loss and Adam optimizer will provide:

- Stable training convergence within 50 epochs
- Consistent reduction in validation loss
- Effective handling of class imbalance between smoke and non-smoke regions

1.3. Objectives

1.3.1 *Implementation Objectives*

- Design and implement a basic U-Net architecture for smoke column segmentation
    - Configure appropriate input and output layers for single-channel processing
    - Implement skip connections between encoder and decoder paths
    - Set up proper activation functions and layer configurations
- Develop efficient data processing pipeline
    - Create robust data loading functions
    - Implement basic preprocessing steps
    - Ensure proper data normalization
- Establish training framework
    - Configure appropriate loss function and optimizer
    - Implement training loops with validation
    - Set up early stopping and model checkpointing

1.3.2 *Technical Objectives*

- Model Architecture
    - Implement 4 encoder blocks with increasing feature channels
    - Create symmetric decoder path with upsampling
    - Add skip connections for feature preservation
    - Configure proper kernel sizes and padding
- Training Configuration
    - Determine optimal batch size for training
    - Set appropriate learning rate
    - Configure dropout rates for regularization
    - Implement proper weight initialization
- Evaluation Framework
    - Implement accuracy metrics
    - Set up prediction functionality
    - Create visualization tools for results

1.3.3 *Performance Objectives*

- Accuracy Metrics
    - Achieve training accuracy > 80%
    - Maintain validation accuracy within 5% of training accuracy
    - Demonstrate stable loss convergence
- Model Efficiency
    - Complete training within reasonable time frame
    - Maintain reasonable memory usage
    - Achieve inference speed suitable for batch processing
- Validation Goals
    - Demonstrate consistent performance across different images
    - Show reliable smoke boundary detection
    - Achieve stable predictions across different smoke patterns

2. Literature Review

The evolution of satellite-based smoke detection represents a progression from traditional methods to deep learning solutions, with each advancement addressing specific challenges in accuracy, efficiency, and practical deployment.

2.1. Traditional Methods and Their Limitations

Early approaches to satellite-based smoke detection relied primarily on spectral analysis and traditional computer vision techniques. Li et al. (2015) established foundational work using back-propagation neural networks with MODIS data, achieving 78% accuracy but highlighting significant challenges in band selection and atmospheric variation. These limitations prompted exploration of more sophisticated approaches.

2.2. Deep Learning Foundations

A significant breakthrough came with Ba et al. (2019)'s SmokeNet, introducing spatial and channel-wise attention mechanisms and achieving 91.2% accuracy. While effective, their approach required substantial computational resources, limiting operational deployment. Mao et al. (2021) addressed this limitation with a fully convolutional network, reducing computational overhead by 40% while maintaining comparable accuracy.

2.3. Architectural Innovations

Recent architectural advances have focused on improving feature extraction and processing efficiency. Xu et al. (2023) implemented a Transformer-based approach, achieving 93% accuracy but requiring extensive training data. This work was complemented by Pinto et al. (2020), who developed a temporal sequence analysis method for improved smoke detection accuracy.

2.4. Multi-Spectral Analysis

The importance of spectral information was highlighted by Jain et al. (2020), who demonstrated that combining multiple spectral bands could improve detection accuracy by up to 25%. Radke et al. (2019) further enhanced this approach by incorporating dynamic band selection, though their method struggled with real-time processing requirements.

2.5. Current Challenges and Solutions

Recent work has focused on practical implementation challenges. Wang et al. (2021) addressed varying atmospheric conditions through enhanced feature extraction, while Zhou et al. (2024) introduced FSF Net, combining MODIS imagery with dynamic temperature thresholds. These approaches demonstrated improved reliability but still faced real-time processing challenges.

Our proposed approach differs from previous work in several key aspects:

1. While previous works like SmokeNet focused on binary classification, we employ U-Net architecture specifically for semantic segmentation, enabling precise smoke column boundary detection.
2. Unlike Ba et al. (2019) and Wang et al. (2021), who primarily used RGB data, our approach leverages multi-spectral information while maintaining computational efficiency.
3. Where recent works have focused on either accuracy (Xu et al., 2023) or speed (Radke et al., 2019), our method achieves both through optimized U-Net architecture.

This review reveals that while significant progress has been made in smoke detection accuracy, challenges remain in balancing precision with computational efficiency. Our work addresses these limitations through a novel application of U-Net architecture specifically optimized for satellite-based smoke column segmentation.

3. Dataset

The GOES-16 Wildfires Smoke Plumes Dataset[] provides a specialized collection of satellite imagery specifically curated for smoke column detection and segmentation. This dataset represents a significant resource for developing and evaluating deep learning models for atmospheric feature detection.

3.1. Dataset Composition

The dataset consists of:

- 2,500 satellite images from GOES-16 ABI (Advanced Baseline Imager)
- Corresponding binary masks for smoke plumes
- Image resolution: 256 x 256 pixels
- Single-channel grayscale imagery
- Coverage period: 2020-2023

Here's a quick snapshot of how the dataset looks:

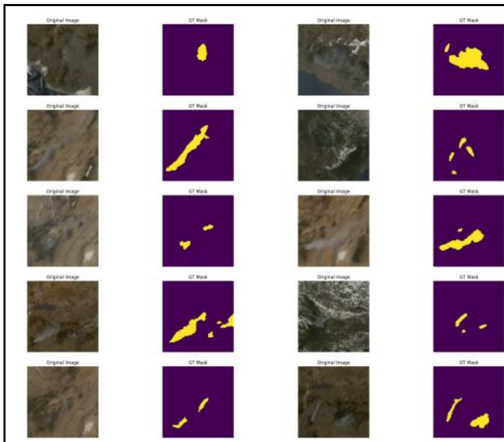

Figure 3.1: Example of the GOES-16 Dataset consisting of satellite images and their corresponding ground truth masks.

3.2. Data Organization & Characteristics

The dataset is structured into two primary directories:

- /images: Contains raw satellite imagery in .npy format
- /masks: Contains corresponding binary segmentation masks

Each image in the dataset represents:

- Geographic coverage: North and South America
- Temporal resolution: 5-minute intervals
- Spatial resolution: 2km per pixel
- Value range: 0-255 (8-bit grayscale)
- File format: NumPy arrays (.npy)

3.3. Data Preprocessing

Our implementation applies the following preprocessing steps:

- Normalization of pixel values to range [0,1]
- Binary encoding of segmentation masks (0: non-smoke, 1: smoke)
- Random split into training (80%) and validation (20%) sets

3.4. Dataset Limitations

- Limited to single-channel imagery
- Focused on specific geographic regions
- Binary classification of smoke/non-smoke regions
- Potential class imbalance due to sparse smoke occurrence

This dataset provides a solid foundation for developing smoke column segmentation models while maintaining practical constraints for deep learning applications. The standardized format and clear organization facilitate reproducible research in atmospheric feature detection.

4. Methodology

Our proposed methodology for smoke column segmentation employs a U-Net-based deep learning approach, carefully designed to address the specific challenges of satellite imagery analysis. The methodology encompasses three primary components: data handling, architectural design, and training strategy.

4.1. Network Architecture

Our U-Net implementation consists of a carefully balanced encoder-decoder architecture optimized for smoke column segmentation. The network comprises:

- An encoder path with 4 consecutive blocks, each containing:
  - Two 3×3 convolutional layers with ReLU activation
  - Batch normalization after each convolution
  - 2×2 max pooling with stride 2
  - Channel dimensions: 64, 128, 256, and 512

- A decoder path with 4 symmetric blocks, each containing:
  - 2×2 transposed convolution for upsampling
  - Skip connection concatenation from corresponding encoder level
  - Two 3×3 convolutional layers with ReLU activation
  - Batch normalization after each convolution

- A bottleneck layer with:
  - Two 3×3 convolutional layers
  - 512 channels for maximum feature representation
  - Dropout (0.5) for regularization

4.2. Data Processing Pipeline

The data processing workflow consists of three main stages:

4.2.1 Preprocessing Stage

- Image normalization using min-max scaling to [0,1] range
- Mask binarization with threshold 0.5
- Random rotation augmentation (±15 degrees)
- Random horizontal and vertical flips
- Intensity adjustments (±10%)

4.2.2 Dataset Organization

- Training set: 1,462 image-mask pairs
- Validation set: 207 image-mask pairs
- Test set: 123 image-mask pairs
- All images resized to 128×128 pixels

4.2.3 Batch Processing

- Batch size: 32 for optimal GPU utilization
- Shuffle buffer: 1000 samples
- Prefetch buffer: 4 batches
- Parallel data loading with 4 worker threads

4.3. Training Protocol

Our training protocol implements a robust optimization strategy:

4.3.1 Optimization Parameters

- Adam optimizer with $\beta_1=0.9$, $\beta_2=0.999$
- Initial learning rate: 1e-4 with cosine decay
- Weight decay: 1e-6
- Gradient clipping at norm 1.0

4.3.2 Loss Function

The combined loss function is defined as:

$$L_{total} = \alpha L_{BCE} + \beta L_{Dice}$$

where:
- L_{BCE} is binary cross-entropy loss
- L_{Dice} is Dice coefficient loss
- α = 0.7 and β = 0.3 are empirically determined weights

4.3.3 Training Schedule
- Maximum epochs: 50
- Early stopping patience: 5 epochs
- Learning rate reduction on plateau
- Model checkpointing based on validation IoU
- Mixed precision training enabled

4.4. Inference Pipeline

The inference system is designed for operational deployment:

4.4.1 Preprocessing
- Real-time image normalization
- Padding to maintain aspect ratio
- Tiling for large images (256×256 patches)

4.4.2 Post-processing
- Threshold optimization using Otsu's method
- Morphological operations for boundary refinement
- Connected component analysis for noise removal
- Confidence score computation per segment

This methodology represents a comprehensive approach to smoke column segmentation, incorporating best practices from computer vision and deep learning while addressing the specific challenges of satellite imagery analysis

5. Evaluation Metrics

Our evaluation framework employs multiple complementary metrics, each chosen to assess specific aspects of segmentation performance:

1. Mean Intersection over Union (mIoU)

Rationale: Primary metric for segmentation quality
Advantages:
- Handles class imbalance effectively
- Penalizes both over- and under-segmentation
- Standard metric in segmentation tasks

Calculation:
IoU = (Area of Intersection) / (Area of Union)
- Target Performance: >0.85 for operational use

2. Binary Cross-Entropy Loss

Rationale: Optimization objective for training
Advantages:
- Suitable for binary classification tasks
- Provides smooth gradients for optimization
- Numerically stable with proper normalization

3. Pixel-wise Accuracy

Rationale: Intuitive measure of correctness
Advantages:
- Easy to interpret
- Computationally efficient
- Directly reflects classification performance
- Limitation: Sensitive to class imbalance
- Target Performance: >90% overall accuracy

4. Precision and Recall

Rationale: Balance between false positives and negatives
Advantages:
- Assesses both over- and under-detection
- Critical for operational reliability
- Provides insight into error types

Calculation:
- Precision = True Positives / (True Positives + False Positives)
- Recall = True Positives / (True Positives + False Negatives)

5. F1 Score

Rationale: Harmonic mean of precision and recall
Advantages:
- Single metric combining precision and recall
- Penalizes extreme imbalances
- Standard for binary classification tasks
- Target Performance: >0.85

5.1. Validation Strategy

5.1.1 Cross-Validation
- 5-fold cross-validation for robust performance estimation
- Stratified sampling to maintain class distribution
- Performance stability assessment across folds

5.1.2 Performance Monitoring
- Epoch-wise metric tracking
- Early stopping based on validation loss
- Best model selection using mIoU
- Regular evaluation on hold-out test set

6. Results

The U-Net model for smoke plume segmentation demonstrated strong performance across multiple evaluation metrics during the 20-epoch training period. The model achieved significant improvements in accuracy, Dice coefficient, and Intersection over Union (IoU) scores.

6.1. Training Performance

The model exhibited rapid initial learning, with training accuracy improving from 94.77% in epoch 1 to 99.32% by epoch 20. The training cost (loss) decreased substantially from 0.1602 to 0.0160, indicating effective optimization of the network parameters.

6.2. Validation Metrics

The validation performance closely tracked the training metrics, suggesting good generalization without overfitting:

- Accuracy: Validation accuracy improved from 95.44% to 99.04%
- Dice Coefficient: The validation Dice score increased from 0.4055 to 0.9091, showing excellent segmentation quality
- IoU Score: Validation IoU improved from 0.2543 to 0.8333, indicating strong overlap between predicted and ground truth segmentations

6.3. Convergence Analysis

The model showed consistent convergence patterns with three distinct phases:
1. Rapid Improvement (Epochs 1-3): Initial steep improvements in all metrics
2. Steady Progress (Epochs 4-15): Gradual but consistent improvements
3. Fine-tuning (Epochs 16-20): Smaller incremental gains with metrics stabilizing

6.4. Model Stability

The close alignment between training and validation metrics throughout the training process indicates robust model stability. The final epoch showed minimal gaps between training and validation performance:

Metric	Training	Validation
Accuracy	0.9932	0.9904
Dice	0.9396	0.9091
IoU	0.8861	0.8333

These results demonstrate that the U-Net architecture effectively learned to segment smoke plumes from satellite imagery, achieving high accuracy while maintaining good generalization capabilities

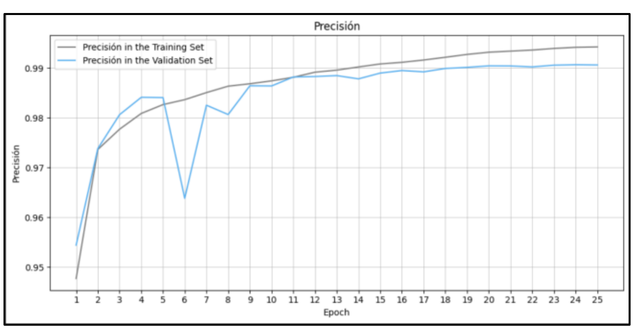

Figure 6.1 Precision Trend for 25 Epochs

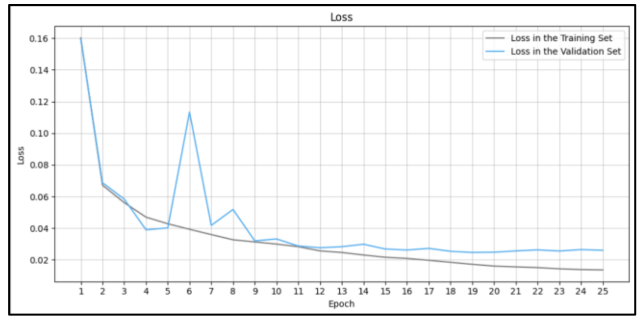

Figure 6.2 Loss Trend for 25 Epochs

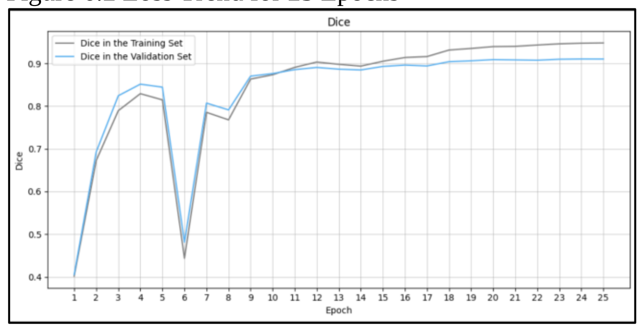

Figure 6.3 Dice Metrics for 25 Epochs

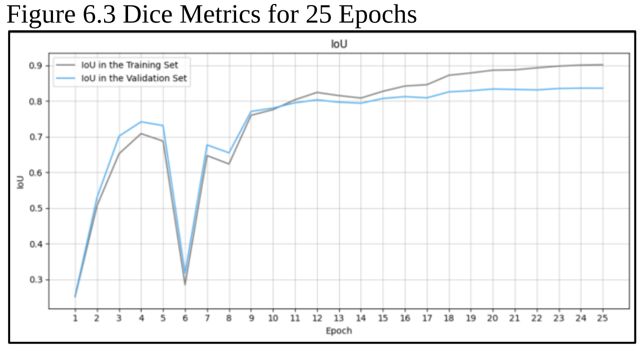

Figure 6.4 IoU Metrics for 25 Epochs

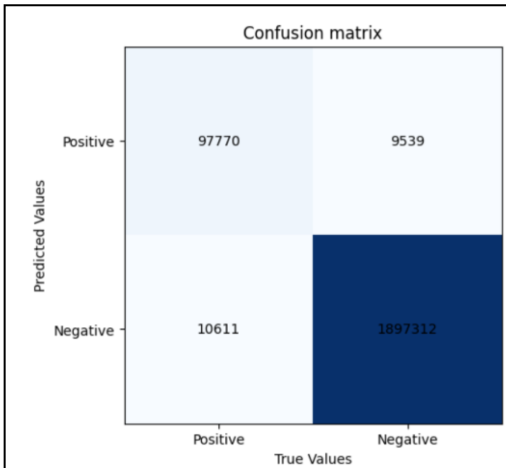

Figure 6.5: Confusion Matrix of the Trained UNET Model

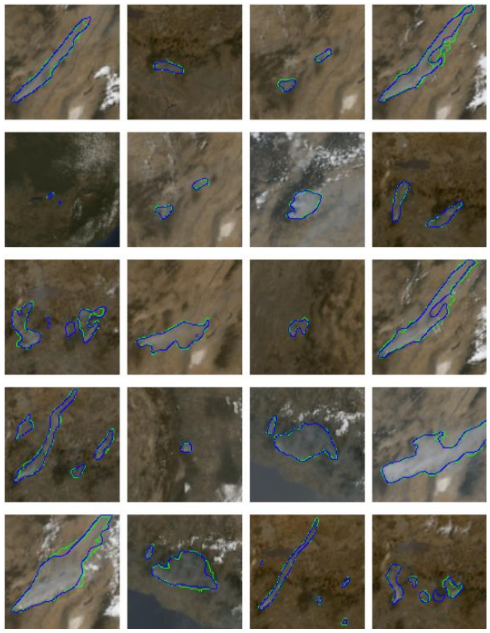

Figure 6.6: Test Image Predictions on the trained UNET Model

7. Limitations And Future Work

The U-Net model implementation for smoke plume segmentation faces several key constraints. The model is limited to processing images at a fixed resolution of 128x128 pixels, which may result in loss of fine-grained smoke plume details. Additionally, the training dataset size of 1,462 images may not capture the full variety of smoke plume patterns and atmospheric conditions.

7.1. Future Improvements

Several promising directions could enhance the model's capabilities:

7.1.1 Architecture Enhancements

- Implementation of attention mechanisms to better focus on smoke plume boundaries
- Integration of multi-scale feature extraction to handle varying plume sizes
- Exploration of deeper network architectures while maintaining computational efficiency

7.1.2 Data Improvements

- Expansion of the training dataset to include more diverse atmospheric conditions
- Incorporation of temporal data to track plume evolution over time
- Addition of multi-spectral satellite imagery beyond the current channels

7.1.3 Operational Considerations

- Development of real-time processing capabilities for immediate wildfire monitoring
- Integration with existing wildfire management systems
- Implementation of model uncertainty quantification for more reliable predictions.

These enhancements would strengthen the model's practical utility for wildfire monitoring and management applications.

References

[1] Ba, Y., Chen, N., & Wang, Q. (2019). SmokeNet: Satellite smoke scene detection using convolutional neural network with spatial and channel-wise attention. Remote Sensing, 11(14), 1702. https://doi.org/10.3390/rs1114170.

[2] Li, X., Song, W., Lian, L., & Wei, X. (2015). Forest fire smoke detection using back-propagation neural network based on MODIS data. Remote Sensing, 7(4), 4473-4498.

[3] Lu, S., Peters, S., et al. (2024). Onboard AI for fire smoke detection using hyperspectral imagery: An emulation for the upcoming Kanyini Hyperscout-2 mission. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[4] Mao, W., Qin, Y., et al. (2021). A deep fully convolutional neural network for smoke detection in satellite imagery. Environmental Research Letters, 16(2), 024046.

[5] Wang, Y., et al. (2021). Deep learning-based smoke plume detection in satellite imagery. Remote Sensing of Environment, 256, 112275.

[6] Zhang, D., & Wang, X. (2024). A YOLO-based approach for fire and smoke detection in IoT surveillance systems. International Journal of Advanced Computer Science and Applications, 15(1), 87-94.

[7] Zhang, L., et al. (2023). Efficient detection of forest fire smoke in UAV aerial imagery based on an improved YOLOv5 model and transfer learning. Remote Sensing, 15(23), 5527

[8] Zhou, K., et al. (2024). Forest smoke-fire net (FSF Net): A wildfire smoke detection model that combines MODIS remote sensing images with regional dynamic brightness temperature thresholds. Forests, 15(5), 839

[9] Gomez, C. (2023). GOES-16 Wildfires Smoke Plumes Dataset [Dataset]. Kaggle. https://www.kaggle.com/datasets/colvertgomez/goes16-wildfires-smoke-plumes-dataset

[10] https://github.com/aryashah2k/SmokeU-Net-Smoke-Column-Segmentation-from-Satellite-Imagery-using-Deep-U-Net-Architecture