Extended Abstract

A Comprehensive Framework for the Automated Optical Gas Imaging of Fugitive Emissions: Integrating Environmentally Adaptive Deep Learning, Novel Data Augmentation, and Hardware-Aligned Spatio-Temporal Morphology for Continuous Monitoring

Department of Computer Science and Engineering

Rajiv Gandhi Institute of Petroleum Technology Jais,

Amethi, Uttar Pradesh, India

Name of Student: Hritu Raj Roll No.: 21CS0003

Email: hritur@rgipt.ac.in

Degree for which submitted: Doctor of Philosophy (Ph.D.)

Dr. Gargi Srivastava

Thesis Supervisor

Overview and Scope of the Thesis

The mitigation of fugitive methane emissions from the oil and gas industry is a critical global objective for ensuring regulatory compliance and combating climate change. While Optical Gas Imaging (OGI) is the premier industry standard for visualizing otherwise invisible hydrocarbon leaks by capturing infrared absorption governed by the Beer-Lambert law, its traditional reliance on manual, intermittent inspections severely limits its effectiveness. Manual OGI drives up labor costs, suffers from human subjectivity, and cannot provide the continuous monitoring required for modern industrial operations. The overarching scope of this thesis is to transition OGI from a manual inspection tool to a continuous, fully automated, and edge-deployable monitoring system utilizing advanced deep learning and computer vision techniques.

The research systematically addresses the fundamental bottlenecks of automated OGI: distinguishing amorphous, fuzzy gas plumes from complex structural backgrounds, adapting to dynamic industrial confounders (e.g., steam, moving clouds, and thermal reflections) that cause false-positive fatigue, and executing heavily parameterized instance segmentation models on resource-constrained edge devices (SWaP-C constraints). By progressing sequentially from foundational binary image classification to feature-augmented instance segmentation, and culminating in an Environmentally Adaptive Human-on-the-Loop (HOTL) framework with a hardware-aligned, spatio-temporal morphological pipeline, this thesis establishes a robust, self-improving blueprint for next-generation Leak Detection and Repair (LDAR) programs.

Chapter 1: Introduction

This chapter provides a comprehensive overview of the global energy landscape, highlighting the dual role of natural gas as both a critical transitional energy source (accounting for roughly 23% of global energy consumption) and a potent greenhouse gas with a global warming potential significantly higher than CO_2 (over 80x over a 20-year period). It establishes the motivation for continuous emission monitoring by quantifying the severe environmental and economic consequences of undetected leaks in aging pipeline infrastructure. The limitations of current manual OGI inspections are critically evaluated, identifying the urgent need for automated computer vision frameworks that can operate reliably and continuously in highly complex, dynamic industrial environments without constant human oversight.

Chapter 2: Fundamentals and Literature Review

This chapter presents a thorough literature survey on gas detection technologies, transitioning from active methods like Tunable Diode Laser Absorption Spectroscopy (TDLAS) and Differential Absorption LiDAR (DIAL) to passive thermal imaging. It details the underlying physics of OGI, governed by the Beer-Lambert law ($A = \epsilon \cdot c \cdot d$) and the differential intensity of infrared radiation ($I = I_0 \cdot e^{-A}$) caused by hydrocarbon absorption. The review critically examines recent advancements in deep learning for gas detection, identifying three major research gaps: (1) the brittleness of static CNN models against site-specific thermal confounders, (2) the loss of high-frequency spatial details and amorphous fluid boundaries during convolutional down-sampling, and (3) the computationally prohibitive nature of running standard segmentation models (like Mask R-CNN) natively on embedded edge hardware without severe latency.

Chapter 3: Automated Optical Gas Imaging for Unattended Gas Leak Detection

To establish the baseline feasibility of replacing manual walk-downs with automated computer vision, this chapter focuses on foundational binary image classification (Leak vs. No-Leak). A custom 2D Convolutional Neural Network (CNN) consisting of optimized Conv-Pool structures, batch normalization, and dropout regularization was developed. To isolate the dynamic gas plumes from static background infrastructure, statistical background subtraction techniques—Gaussian Mixture Model (GMM) and Mixture of Gaussians (MOG)—were applied using a rolling average over 150 frames. The model was trained and evaluated on an extensive dataset comprising 32,000 FLIR frames and 837 Kaggle smoke images, categorized across 6 leak sizes (from barely visible $0 - 100\ g/h$ to very high $1601 +\ g/h$) and 5 imaging distances (4.2 m to 18.3 m). The proposed GMM-CNN framework achieved an impressive 99% accuracy for medium-to-large leaks at close ranges and 94% for small leaks. Crucially, GMM consistently outperformed MOG by up to 6.2% at extended distances (18.3 m), proving that machine learning can reliably identify gas emissions without human intervention.

Chapter 4: High-Fidelity Gas Plume Segmentation via Bit Plane Slicing Augmentation

Moving beyond simple binary classification, this chapter tackles the complex task of precise spatial delineation (instance segmentation) of gas plumes. Because gas plumes often exhibit exceptionally low thermal contrast against their backgrounds, a novel data augmentation strategy utilizing Bit Plane Slicing (BPS) was introduced. By decomposing 8-bit thermal frames and applying an XOR operation between bit planes 4 and 5 (Bit 4 $\oplus$ Bit 5), subtle thermal gradients and motion vectors were computationally amplified, mimicking high-sensitivity OGI modes. Evaluating combinations, (Bit 4 $\oplus$ Bit 5) best emphasized fine fluid transitions. Integrating this enriched 2,000-image dataset with a Mask R-CNN (ResNet-101 FPN) architecture yielded an F1-Score of 98.7%, a Dice Coefficient of 0.965, and a Mean Average Precision (mAP) of 0.853. This analytical augmentation strategy significantly outperformed traditional image enhancement techniques like CLAHE and Gamma correction (F1-score 96.3%), providing crisp, binary-like boundary representations for the deep feature heads.

Chapter 5: Environmentally Adaptive Tiered Human-on-the-Loop (HOTL) Framework

Static deep learning models inevitably suffer from false-positive fatigue in real-world deployments due to site-specific confounders like steam vents, exhaust, and solar reflections. This chapter introduces an Environmentally Adaptive Human-on-the-Loop (HOTL) framework to solve this critical industrial challenge. The operational core is a novel 4-stage tiered pipeline (CNN Screening → SSD Localization → On-Edge Mask-RF → On-Server Mask-RF) that smartly balances computational loads. A new Mask-Refinement (Mask-RF) model was introduced, utilizing a feedback loop architecture where final logits are derived from averaging initial coarse predictions and original RoI features ($L_{final} = \frac{L_{initial}+L_{refined}}{2}$). This improved plume delineation Average Precision (AP) to 67.18%, a relative 7.7% boost over standard models, and reduced loss by 15.8%. Crucially, the system autonomously flags low-confidence predictions for asynchronous human review, curating a "hard-case" dataset. This exception-based learning loop allows the models to undergo adaptive retraining, progressively mastering the specific visual signature of its industrial environment and minimizing alert fatigue.

Chapter 6: Real-Time Edge-Aligned Pruning and Spatio-Temporal Watershed Segmentation

This chapter represents the culmination of the research, bridging the gap between deep learning accuracy and edge hardware limitations (SWaP-C constraints). A hardware-aligned architecture, HA-TaMRCNN, was developed using 1st-order Taylor Expansion saliency pruning evaluated via $S(G_i) = \sum | W \frac{\partial L}{\partial W} |$. By enforcing strict Modulo-16 channel constraints ($C_{out} \pmod{16} \equiv 0$), the model bypassed cuDNN padding penalties to maximize NVIDIA Tensor Core efficiency, followed by Foreground-Masked Knowledge Distillation to recover capacity. To capture the fuzzy, chaotic boundaries of the plume without stalling the CPU, a Temporal Variance-Guided (TVG) Watershed algorithm was introduced. By anchoring seeds based on spatial variance ($T_{fg} = \mu_{std} + 1.5\sigma_{std}$) and executing local Minkowski morphology strictly in GPU VRAM (via CuPy) over a Dynamic Bounding Box Union ($ROI = BBox_{AI} \cup BBox_{\sigma>T}$), PCIe bottlenecks were eliminated. Deployed on an NVIDIA Jetson Orin edge device, the pipeline cut FLOPs from 275.4 to 122.5 GFLOPs (55.5%) and VRAM consumption by 48.5%. The system achieved a real-time throughput of 58.4 FPS (using TensorRT FP16) while elevating boundary accuracy (AP_{75}) by 11.9% to 78.40% and drastically reducing the global false-positive rate from 12.80% to an unprecedented 3.20%.

Chapter 7: Conclusion and Future Scope

This chapter summarizes the primary contributions of the thesis, highlighting how the progression from basic CNN classification to hardware-software co-designed edge segmentation provides a complete, scalable, and self-improving solution for automated LDAR programs. The hybrid models successfully bridge the gap between qualitative visualization and precise, automated analytical capabilities. Future research directions are outlined, including the integration of stereo-vision depth estimation utilizing epipolar geometry with the 2D output contours of the TVG-Watershed algorithm to transition toward 3D volumetric mass emission modeling in physical space. Furthermore, adapting the lightweight pipeline for deployment on Unmanned Aerial Vehicles (UAVs) combined with multi-sensor fusion (acoustic and pressure data) holds significant potential for large-scale, continuous mobile facility monitoring and digital twin integration.

List of Publications:

Raj, H., & Srivastava, G. (2026). A novel data augmentation strategy for gas leak detection and segmentation using Mask R-CNN and bit plane slicing in chemical process environments. Computers and Chemical Engineering, 204, 109407. (**Published**)

Raj, H., & Srivastava, G. (2025). An OGI Based Unattended Natural Gas Leak Detection System by Utilizing the Power of Machine Learning and Computer Vision. In Proceedings of the 14th International Conference on Pattern Recognition Applications and Methods (ICPRAM), 823-830. (**Published**)

Raj, H., & Srivastava, G. A survey on recent trends and approaches for detecting leakage of invisible natural gas based on thermal imaging techniques. Submitted as a book chapter in Next-Generation Computing: Applied Innovations in IT from Research to Practice (Bentham Science). (**Accepted**)

Raj, H., & Srivastava, G. Enhancing Process Safety in the Oil and Gas Industry: An Environmentally Adaptive Tiered Human-on-the-Loop Framework for Continuous Methane Leak Monitoring, is submitted in Machine Vision and Applications. (**Submitted**)

Raj, H., & Srivastava, G. A Real-Time Hybrid Deep Learning and Spatio-Temporal Morphological Pipeline for Natural Gas Plume Segmentation using HA-TaMRCNN and TVG-Watershed, is under review in Journal of Real-Time Image Processing. (**Under Review**)