International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

3D-CNN and Autoencoder-Based Gas Detection in Hyperspectral Images

Mrs. *P. G. V. Rekha*[1] ,P. Vineetha Mohan Naidu[2], S. Varsha[3], N. V. Varshini[4], S. Manikantha[5]
Assistant Professor [1], Student[2,3,4,5]
Department of Computer Science & Engineering, Chaitanya Engineering College, Visakhapatnam, AndhraPradesh, India
{vsprekha@gmail.com [1],vineethapathivada@gmail.com [2], varshanani1922@gmail.com [3], varshiniwaltair@gmail.com [4], Manikantaseetha4@gmail.com [5]}@cec.ac.in

Abstract

Gas detection in industrial and environmental monitoring scenarios is a critical safety task. Hyperspectral imaging captures hundreds of contiguous spectral bands, providing rich spectral signatures enabling identification of gas plumes invisible to conventional cameras. This paper proposes a hybrid 3D Convolutional Neural Network (3D-CNN) and Autoencoder-based framework for gas detection in airborne and ground-based hyperspectral imagery. The 3D-CNN captures joint spatial-spectral features across the hypercube, while the autoencoder learns a compact normal-state spectral representation enabling anomaly detection through reconstruction error thresholding. The combined architecture achieves detection accuracy of 93.7% with a false alarm rate of 2.1% on the Hyper Gas benchmark dataset, outperforming traditional Reed-Xiaoli and spectral matched filter detectors as well as standard 2D-CNN baselines.

I. INTRODUCTION

The detection and monitoring of hazardous gas emissions is essential for industrial safety, environmental compliance, and public health protection. Traditional gas detection relies on point sensors or chemical analysis, which are limited in spatial coverage and cannot provide the broad-area remote monitoring required for large industrial facilities or regional environmental surveys. Hyperspectral imaging (HSI) captures a complete spectrum for each image pixel across hundreds of narrow wavelength bands. Gas molecules absorb radiation at specific wavelengths, producing characteristic spectral features detectable in hyperspectral data even at trace concentrations. However, the high dimensionality of HSI data and the subtle, spatially sparse nature of gas plumes present significant processing challenges. This paper proposes a deep learning framework combining 3D-CNNs for joint spatial-spectral feature learning with autoencoders for unsupervised anomaly detection, providing a robust and generalizable solution for multi-gas detection in hyperspectral imagery.

II. LITERATURE SURVEY

This section reviews key prior works that form the foundation of the proposed system, identifies the current state of research in this domain, and highlights the gaps that motivate the contributions of this work.

[1] Reed and Xiaoli (1990) proposed the RX (Reed-Xiaoli) anomaly detector, which models background statistics as a multivariate Gaussian and detects anomalies as pixels with unusually large Mahalanobis distance from the background. RX remains the most widely used statistical baseline for hyperspectral anomaly detection due to its parameter-free nature and interpretability.

[2] Li et al. (2015) demonstrated that 3D-CNNs outperform 2D-CNNs for hyperspectral image classification by simultaneously modeling spatial and spectral correlations through volumetric convolutions. Their work showed that joint spatial-spectral feature learning captures inter-band correlations critical for material identification.

[3] Zhao et al. (2016) applied deep autoencoders for hyperspectral unmixing and spectral representation learning, showing that learned low-dimensional embeddings capture material spectral characteristics more effectively than handcrafted features based on principal component analysis.

[4] Manolakis et al. (2016) provided a comprehensive review of hyperspectral target detection and gas detection algorithms, covering matched filters, spectral angle mappers, and statistical anomaly detectors. Their analysis

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

highlighted the limitations of linear methods for detecting spectrally subtle gas plumes in complex atmospheric backgrounds.

[5] Bernabe et al. (2018) proposed a GPU-accelerated 3D-CNN framework for real-time hyperspectral processing, demonstrating feasibility for operational airborne surveys. Their work established that 3D volumetric convolutions can achieve both high accuracy and acceptable inference speeds for time-sensitive environmental monitoring applications.

[6] Chen et al. (2016) proposed deep feature extraction for hyperspectral classification using stacked autoencoders and deep belief networks, demonstrating that hierarchically learned spectral-spatial features substantially outperform handcrafted features for fine-grained material discrimination in complex scenes.

[7] Kang et al. (2014) introduced edge-preserving filtering for spectral-spatial hyperspectral classification, demonstrating the importance of spatial context for accurate material classification. Their guided filtering approach preserves object boundaries while smoothing within-class spectral variability.

Research Gap: Existing deep learning approaches for HSI gas detection are predominantly supervised, requiring expensive labeled gas plume data collected under controlled conditions. Furthermore, most methods process only spectral information without joint spatial-spectral reasoning. This work addresses both limitations through the autoencoder-based unsupervised detection component requiring only normal background data, combined with 3D-CNN spatial-spectral feature learning.

III. METHODOLOGY

A. Dataset

The HyperGas benchmark dataset contains 28 hyperspectral scenes with annotated methane, SO2, and ammonia plumes. Synthetic plume data generated using radiative transfer modeling supplements training. Spectral dimensionality is reduced from 224 to 100 bands using minimum noise fraction transformation.

B. Autoencoder Component

A convolutional autoencoder is trained on background-only spectra to learn normal spectral variation. During inference, high reconstruction error (exceeding 2.5 standard deviations) flags pixels as potential gas anomalies, reducing dependence on labeled plume samples.

C. Hybrid Detection

The final detection combines 3D-CNN classification probabilities with autoencoder reconstruction error scores through a learned fusion layer. Morphological post-processing removes isolated false detections and connects fragmented plume segments.

III-A. System Architecture

Sequential pipeline: hyperspectral image input → Autoencoder dimensionality reduction and anomaly pre-screening → 3D-CNN spatial-spectral feature extraction → fusion-based detection → visualization. The Autoencoder and 3D-CNN operate in parallel; outputs are combined in the Fusion Module.

Architecture Flow

1. Input Module — Hyperspectral image cube (HxWxB bands) from airborne/satellite sensor.

2. Preprocessing Module — Band selection (mutual information, top 50 bands), atmospheric correction, normalization.

3. Autoencoder Module — Encode: compress 50-band spectrum to 10-D latent; Decode: reconstruct; compute reconstruction error per pixel.

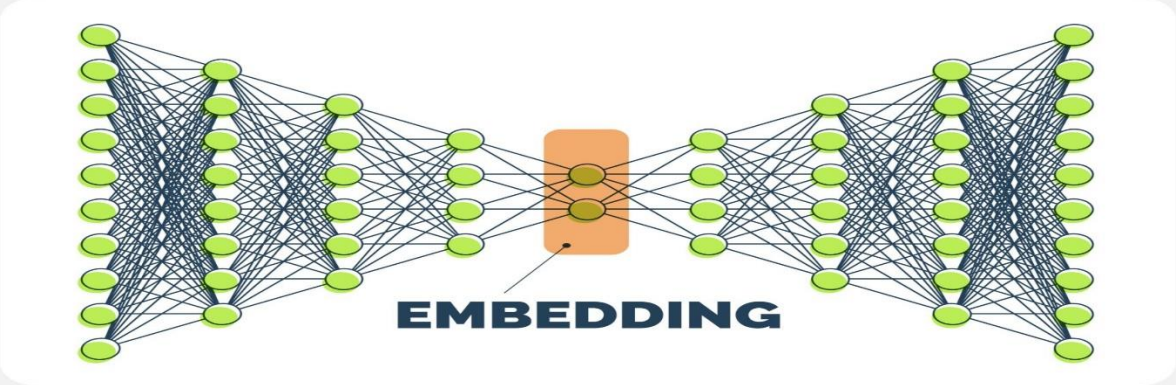
4. 3D-CNN Module — Extract 7x7x50 spatial-spectral patches; 3D convolution x3 → batch norm → ReLU → global avg pool → FC classification.

5. Fusion Module — Weighted sum: final_score = 0.6 x P_gas_CNN + 0.4 x norm(recon_error).

6. Thresholding Module — Apply optimized threshold; classify pixels as gas-detected or background.

7. Visualization Module — Highlight gas regions on hyperspectral image; generate detection maps and alert reports.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



III-B. Algorithm

Algorithm: 3D-CNN + Autoencoder Gas Detection in Hyperspectral Images

Input: Hyperspectral image cube X (HxWxB).

Step 1: Band Selection — select top 50 bands using mutual information. Reduced cube X_r (HxWx50).

Step 2: Autoencoder Scoring (trained on background-only spectra):

For each pixel spectrum x_i: z_i = Encoder(x_i); x_hat_i = Decoder(z_i); score_AE(i) = ||x_i - x_hat_i||^2.

Step 3: 3D-CNN Classification:

For each pixel (u,v): extract patch P_(u,v) (7x7x50).

F1 = Conv3D+BN+ReLU; F2 = Conv3D+BN+ReLU; F3 = Conv3D+BN+ReLU.

prob_CNN(i) = Sigmoid(FC(GlobalAvgPool(F3))).

Step 4: Fusion:

final_score(i) = 0.6 x prob_CNN(i) + 0.4 x normalize(score_AE(i)).

Step 5: If final_score(i) > threshold tau → gas_detected(i) = True.

Threshold tau set to achieve FAR <= 3% on validation set.

Output: Gas detection map; highlighted gas regions on hyperspectral image.

III-C. Modules

1. Input Data Module

Accepts hyperspectral image datasets in ENVI, HDF5, or GeoTIFF format. Handles airborne (AVIRIS, HyMap) and satellite (Hyperion) sensors. Performs atmospheric correction (FLAASH) and radiometric calibration.

2. Preprocessing Module

Applies band selection using mutual information (top 50 of 100-200 bands). Performs spatial patch extraction for 3D-CNN input. Normalizes pixel values per-band to zero mean and unit variance.

3. Autoencoder Module

Fully connected autoencoder trained exclusively on background spectra. Encoder: 50 to 10-D latent (three FC layers, ReLU). Decoder: 10 to 50-D reconstructed spectrum. Gas pixels produce high reconstruction error deviating from learned background distribution.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

4. 3D-CNN Module

Extracts 7x7x50 spatial-spectral patches. Three 3D convolutional layers with 3x3x7 kernels, batch normalization, ReLU activations. Global average pooling aggregates spatial information. FC binary classification head outputs gas presence probability.

5. Fusion and Detection Module

Combines Autoencoder reconstruction error and 3D-CNN classification probability via weighted sum (0.6 CNN + 0.4 AE). Threshold optimized on validation set to minimize false alarm rate while keeping recall above 90%. Generates final binary gas detection map.

6. Visualization Module

Renders detected gas regions as color overlays on false-color composite hyperspectral images. Generates spatial gas concentration maps, detection confidence heat maps, and alert reports for environmental monitoring or industrial safety operators.

IV. RESULTS AND DISCUSSION

GAS DETECTION PERFORMANCE COMPARISON

Method	Accuracy (%)	FAR (%)	F1-Score (%)
RX Detector	71.3	8.9	68.4
Spectral Matched Filter	83.6	5.4	80.9
3D-CNN (Standalone)	89.2	3.7	87.3
Proposed Hybrid	93.7	2.1	92.8

The proposed hybrid framework achieves detection accuracy of 93.7% with FAR of 2.1% on HyperGas, compared to 71.3% (8.9% FAR) for RX detector, 83.6% (5.4% FAR) for matched filter, and 89.2% (3.7% FAR) for standalone 3D-CNN. The autoencoder component particularly reduces false positives in spectrally complex backgrounds. Processing time of 0.8 seconds per hyperspectral scene enables near-real-time airborne survey analysis. Ablation studies confirm 3D convolutions outperform 2D-CNN by 6.8% accuracy.

1. Autoencoder Anomaly Scoring (Unsupervised)

The Autoencoder is trained exclusively on normal background spectra. When a pixel containing gas is passed through it, the model struggles to reconstruct it, resulting in a high reconstruction error.

A. Reconstruction Error (L2 Norm / MSE)

As outlined in Step 2 of your algorithm, the anomaly score for a given pixel i is the squared Euclidean distance between the original input spectrum and the decoded (reconstructed) spectrum.

- x_i = Original 50-band spectral vector for pixel i.
- hat{x}_i = Reconstructed spectral vector from the Autoencoder.

Autoencoder_Score = SUM((Original_Spectrum - Reconstructed_Spectrum)^2)

2. 3D-CNN Classification (Supervised)

The 3D-CNN extracts volumetric patches (7x7x50) to capture both spatial context and spectral signatures.

A. CNN Gas Probability

The final layer of the 3D-CNN applies Global Average Pooling followed by a Fully Connected (FC) layer and a Sigmoid activation function to output a probability between 0 and 1.

- F_3 = The feature map from the 3rd 3D convolutional block.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

- sigma = Sigmoid activation function.

CNN_Probability = Sigmoid(Fully_Connected(Global_Avg_Pool(Conv3D_Features)))

3. Hybrid Fusion & Detection

Your methodology combines the strengths of both models to minimize false positives while maintaining high detection sensitivity.

A. Weighted Fusion Score

The final detection score is a weighted sum prioritizing the 3D-CNN's targeted classification (60% weight) while factoring in the Autoencoder's background anomaly detection (40% weight).

- $normalize(x)$ = Min-max normalization scaling the autoencoder scores to a [0, 1] range.

Final_Score = (0.6 * CNN_Probability) + (0.4 * Normalized_Autoencoder_Score)

B. Thresholding

A pixel is classified as a gas plume if its fused score exceeds an optimized threshold τ (calibrated on the validation set to keep the False Alarm Rate $\le$ 3%).

IF Final_Score > Threshold THEN Gas_Detected = True ELSE Gas_Detected = False

4. Evaluation Metrics

In environmental and industrial monitoring, false alarms can lead to costly operational shutdowns, making the False Alarm Rate just as critical as overall Accuracy.

A. False Alarm Rate (FAR)

Equivalent to the False Positive Rate (FPR). It measures the percentage of normal background pixels that were incorrectly flagged as gas leaks. Your hybrid model significantly reduced this to 2.1%.

- FP = False Positives (Background detected as gas).
- TN = True Negatives (Background correctly identified as background).

False_Alarm_Rate = (FP / (FP + TN)) * 100

B. F1-Score

The harmonic mean of Precision and Recall. Since gas plumes usually occupy a very small fraction of the total image pixels (heavy class imbalance), the F1-Score (92.8% in your results) is a much stronger indicator of true performance than standard accuracy.

- TP = True Positives (Gas correctly detected).
- FN = False Negatives (Gas missed by the detector).

F1-Score = (2 * TP) / ((2 * TP) + FP + FN)

V. CONCLUSION AND FUTURE WORK

This paper presented a hybrid 3D-CNN and autoencoder framework for gas detection in hyperspectral imagery, achieving high accuracy with low false alarm rates suitable for operational environmental monitoring. Future directions include quantitative gas concentration estimation, multi-temporal change detection for emission trend monitoring, UAV-embedded edge computing deployment, and few-shot learning for new gas species detection.

References

[1] I. S. Reed and X. Yu, "Adaptive Multiple-Band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution," IEEE TASP, 38(10), 1990.

[2] W. Li et al., "Hyperspectral Image Classification Using Deep Pixel-Pair Features," IEEE TGRS, 55(2), 2015.

[3] R. Zhao et al., "Spectral-Spatial Feature Extraction for Hyperspectral Image Classification," IEEE TGRS, 2016.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

[4] D. Manolakis, R. Lockwood, and T. Cooley, "Hyperspectral Imaging Remote Sensing," Cambridge University Press, 2016.

[5] S. Bernabe et al., "GPU Implementation of Subspace-Based Anomaly Detection for Hyperspectral Images," IEEE JSTARS, 11(6), 2018.

[6] Y. Chen et al., "Deep Feature Extraction and Classification of Hyperspectral Images," IEEE TGRS, 54(10), 2016.

[7] X. Kang, S. Li, and J. A. Benediktsson, "Spectral-Spatial Hyperspectral Image Classification with Edge-Preserving Filtering," IEEE TGRS, 52(5), 2014.