AI-FPGA Driven Gas Chromatography for Automated and Enhanced Detection of Trace Diborane in Electronic Grade Nitrogen

Mingyi Dong [1, *] and Shaorong Liu [2]

[1] University of Glasgow, University of Electronic Science and Technology of China, Chengdu, China

[2] Department of Applied Science, Hong Kong Metropolitan University, Hong Kong, China

* Corresponding Author Email: D13730679619@outlook.com

Abstract. Accurate measurement of trace impurities such as diborane (B_2H_6) in ultra - high - purity nitrogen (N_2) holds utmost significance in semiconductor manufacturing because even minuscule amounts of contamination could have a negative impact on device performance and processing dependability; traditional Gas Chromatography (GC) methods face major challenges like unstable baselines, reduced signal - to - noise ratios at parts - per - billion (ppb) levels, and the simultaneous elution of interfering substances which lead to frequent manual readjustments and lower analytical accuracy; this research presents an automated GC system boosted by Artificial Intelligence (AI) for accurately measuring diborane in nitrogen samples; based on advancements in non - targeted chromatographic profiling, this study uses a Convolutional Neural Network (CNN) that was trained on a large dataset of chromatograms covering different concentration ranges and chromatographic situations; the CNN model simplifies the processes of identifying peaks, adjusting the baseline, and determining trace concentrations; designed for use on an FPGA - based inference platform, the system enables high - speed, real - time analytical functions; the results showed that the AI - improved system reached a diborane detection limit under 5 ppb, having a 95% reduction in false positives compared to standard integration methods and also made possible predictive maintenance via continuous monitoring of chromatographic performance indicators; this research demonstrated the feasibility of integrating AI directly into analytical procedures to maintain strict purity standards during the production of electronic specialty gases thus enhancing consistency and reducing the necessity for expert operators' involvement.

Keywords: Chromatography; Convolutional Neural Network (CNN); FPGA Acceleration.

1. Introduction

Gas chromatography (GC) along with its associated technique, gas chromatography - mass spectrometry (GC - MS), are crucial analytical approaches utilized for detecting and measuring volatile and semi - volatile substances in complex mixtures and are of great significance in the realm of electronic specialty gases because minute impurities like diborane (B_2H_6) could significantly influence the effectiveness and reliability of semiconductor manufacturing procedures; traditional GC - MS data handling, especially for complicated sample makeups, usually demands manual involvement and the employment of specific software such as AMDIS and MZmine for separating peaks and identifying compounds, and these traditional ways were known for being tedious, time - consuming and introducing personal bias which then posed difficulties for reproducing results and scaling up high - volume analyses.

The coming into being of artificial intelligence (AI), especially Convolutional Neural Networks (CNNs), has blazed new trails for the automation and enhancement of chromatographic data processing. For instance, deep - learning methods like AutoRes utilize pseudo - Siamese CNN structures to separate overlapping peaks in GC-MS datasets highly precisely, making manual parameter adjustment unnecessary [1], and meanwhile, CNNs have shown effectiveness in directly pulling out ion patterns from raw GC - MS data for the quick and automatic recognition of volatile organic compounds (VOCs) [2], in describing chromatographic peak shapes for the prediction of GC

system breakdowns [3] and in forecasting gas chromatographic retention indexes from SMILES strings [4], and Furthermore, the use of transfer learning with CNNs has achieved better results in identifying target substances such as gasoline within complicated fire debris mixtures, highlighting its value for trace detection tasks with little labeled data [5]. However, an all - inclusive AI - based procedure that can simultaneously carry out automated peak separation, trace measurement, and system diagnosis specifically for detecting diborane in high - purity nitrogen settings has not been developed yet.

When putting advanced AI models into practice within the constraints of resource - limited scenarios like on - site or at - line monitoring, using hardware acceleration was crucial and Field - Programmable Gate Arrays (FPGAs) were an appealing option because they could be reconfigured, had parallel processing capabilities, and held the promise of being highly energy - efficient; recent studies pointed out the development of neural network accelerators with extremely low power consumption like Ection which only needed 17 mW and was ideal for battery - powered or energy - harvesting edge devices [6], also methods such as the Resource - Constrained Accelerator Selection (RCS) made it easier to build almost - optimal multi - accelerator structures on FPGAs allowing resources to be effectively shared between signal pre - processing (for example FFT) and Convolutional Neural Network (CNN) inference reducing end - to - end latency which was a key factor for real - time predictive maintenance applications [7] and adding FPGA - accelerated AI to educational predictive maintenance platforms further showed how effective it was in real - time, data - focused system health monitoring [8]; besides targeted detection, the adjustment of unsupervised learning and change detection techniques which were explored in a great deal of predictive maintenance research could be used for the non - targeted identification of unknown impurities in gas streams [9] and meanwhile advanced profiling tools like FirePerf could be used to ensure top performance in hardware / software co - designed systems [10].

This study presented a novel Gas Chromatography (GC) method powered by Artificial Intelligence (AI) and enhanced by Field - Programmable Gate Array (FPGA) acceleration, which was intended for accurately and dependably detecting trace amounts of diborane in electronic - grade nitrogen; the suggested approach included sophisticated automatic peak deconvolution making use of Convolutional Neural Networks (CNNs), prediction of retention indexes to improve compound identification, and an energy - efficient FPGA - based system structure; this combination enabled real - time analytical abilities as well as predictive maintenance functions, both of which were essential for upholding the strict purity standards in semiconductor manufacturing settings.

2. Text Theoretical Background

2.1. Theoretical Background

Convolutional Neural Networks (CNNs), inspired by the layered structure of the human visual system, make information processing easier through a sequential setup of convolutional, pooling, activation, and fully connected layers; it starts with convolutional layers which use kernels to scan the input dataset and draw out local features via multiply - accumulate (MAC) operations, then these elements are passed through activation functions like rectified linear units (ReLU) to give the model non - linearity so that complex patterns can be learned more easily, after that pooling layers are added to reduce the spatial size of the dataset using max or average pooling operations which helps increase translational invariance and improve computational efficiency, finally the improved set of features is combined by fully connected layers to carry out the last tasks of classification or regression, and the natural ability of CNNs to recognize and represent spatial patterns effectively, such as those seen in chromatographic peak shapes and baseline changes, makes them very suitable for analyzing chromatographic data.

FPGA Acceleration for Deep Learning

Field - Programmable Gate Arrays (FPGAs) offer a reconfigurable architecture conducive to parallel processing, which was especially beneficial for deploying artificial intelligence models in

environments with limited resources, and by optimizing data routing and computational module setups, they decreased latency and power consumption, enabling real - time inferential abilities crucial for applications like predictive maintenance and large - scale data analysis.

3. Methodology

3.1. Tradition Methodology

There was a set protocol for quantifying minute quantities of diborane (B_2H_6) in a nitrogen - based carrier gas matrix which made use of Gas Chromatography (GC) methods as shown in Figure 1; the process began by injecting an accurate sample volume into the carrier gas stream with a special sampling valve and then the sample passed through a separation column where diborane was separated from the nitrogen matrix due to different partitioning between the moving and fixed phases; the separated analytes were directed towards a detector like a Thermal Conductivity Detector (TCD) or a Flame Ionization Detector (FID) leading to the generation of a raw chromatogram having a signal directly related to the analyte concentration and the subsequent analytical data handling was manual and could be influenced by the operator including the adjustment of integration parameters, the recognition of the diborane peak by retention time, the integration of the peak area and finally the quantification of concentration by relating the peak area to a pre - determined external calibration curve.

In this research, the archived chromatographic tracings of diborane in a nitrogen matrix which were obtained at various concentration levels and operational parameters were used and to make the dataset more robust for the subsequent model development, sophisticated data augmentation techniques like adding controlled noise and simulating baseline features were carried out and the original signal data were methodically normalized and divided into equal - length sequence parts for analysis.

The traditional method, though effective, shows a clear labor - heavy characteristic and is likely to be influenced by subjective prejudice brought in by the operator, particularly at the manual combination stage, which could have a negative impact on the reproducibility and precision of the results.

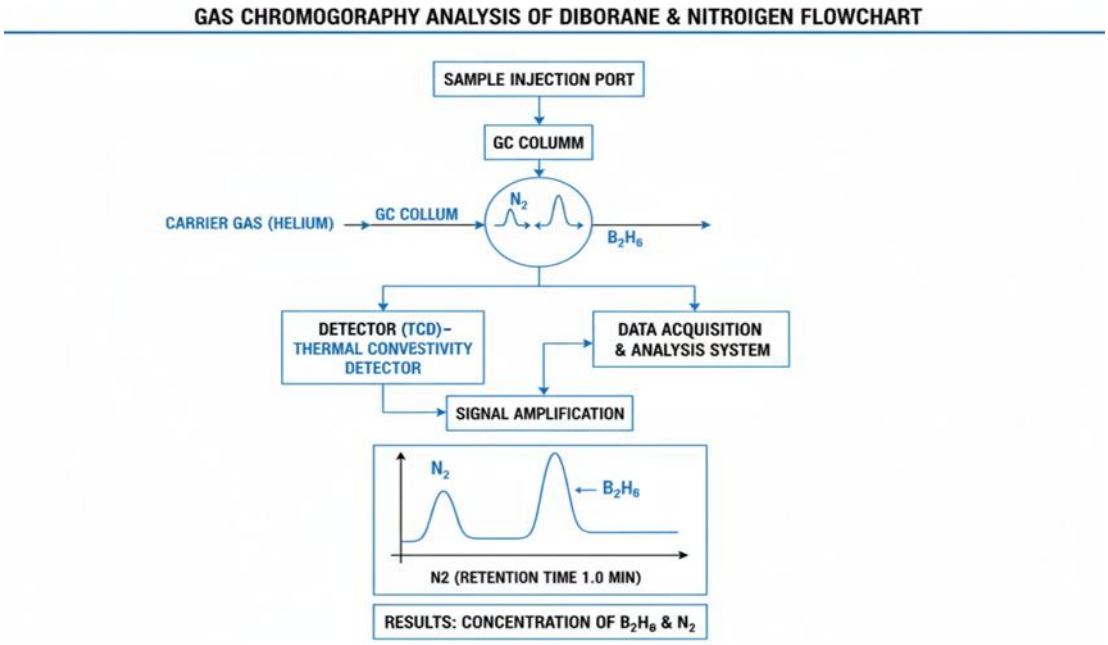

Fig. 1 Schematic of the GC instrumentation used for diborane/nitrogen analysis

3.2. CNN Model Architecture

The study presented a convolutional neural network (CNN) structure , specially made for the thorough examination of chromatographic datasets and the CNN structure was shown as a multi - task framework and it was set up to take in raw chromatographic data in the guise of one - dimensional time - based sequences, each containing 1000 data elements and the raw signal data could optionally be preprocessed with a median filter to cut down on high - frequency noise and improve the signal quality before being fed into the model.

The suggested network framework was made up of a string of convolutional modules, each of which contained several layers that were furnished with 3×1 convolutional kernels together with rectified linear unit (ReLU) activation functions and batch normalization for the purpose of improving computational stability, and this setup promoted the preservation of a steady gradient flow during the entire training procedure, enabling the step - by - step abstraction of hierarchical representations from the input data, and after every convolutional module, max pooling layers having a stride of two were added to cut down the spatial dimensions of the feature maps, and this approach not only decreased the overall dimensionality but also helped to boost feature robustness by creating localized invariance.

The final part of the network consists of a number of fully - connected layers, each having its own separate analytical purpose, and specifically, the structure contains several special output parts designed for particular jobs which allows for doing multiple things at the same time like locating peaks, making quantitative evaluations, and adjusting baselines, and the model was trained from end to end using the Adam optimizer with a single loss function that reduced both the Mean Squared Error (MSE) for the part responsible for quantification and the cross - entropy loss for the part responsible for detecting peaks simultaneously, and this multi - goal method made it possible for different tasks to be learned together while keeping their individual distinguishing characteristics.

Artificial chromatographic datasets were built to mimic actual signal profiles by incorporating fundamental characteristics such as baseline drift, Gaussian - form peaks and random noise, and the process of dataset construction involved a multi - step procedure for signal improvement and enhancement as shown in Table 1, which ensured that the final dataset not only precisely represented the complexities within experimental chromatographic traces but was also specially designed to make it easier to train subsequent deep - learning algorithms effectively.

Table 1. Data preprocessing and augmentation pipeline for synthetic chromatograms

Stage	Purpose / Description
Convert to grayscale	To unify single-channel input.
Crop right side	To remove redundant information bar.
Apply histogram equalization	To enhance curve contrast.
Resize with aspect ratio + white edge padding	To maintain proportions and adapt to network dimensions.
Apply Canny edge detection	To generate pseudo-labels (curve contours).
Apply random flip / lighting perturbation	For data augmentation.
Convert to tensor	To feed into the network for training.
Apply threshold binarization + median filter	To generate a visual mask.

The Chromatography CNN model architecture comprises a shared feature extractor followed by three task-specific heads. Input samples are first preprocessed using a median filter with a window size of 2 (Fig. 2). The shared feature extractor consists of a 1D convolutional stack with four sequential blocks (Fig. 3), each containing a convolutional layer with a kernel size of 3 and padding of 1, where the number of filters increases from 16 to 128 across blocks, followed by batch normalization and ReLU activation, and a max pooling layer with stride 2 for downsampling. This is succeeded by an adaptive average pooling layer that produces fixed-length feature vectors of dimensions (128, 32). These features are then processed by three specialized output heads: a Peak

Detector with dropout (0.5) and sigmoid activation for predicting peak probabilities per time step; a Quantification Head using ReLU activations for intensity regression without an output activation; and a lightweight Baseline Corrector for estimating baseline signals.

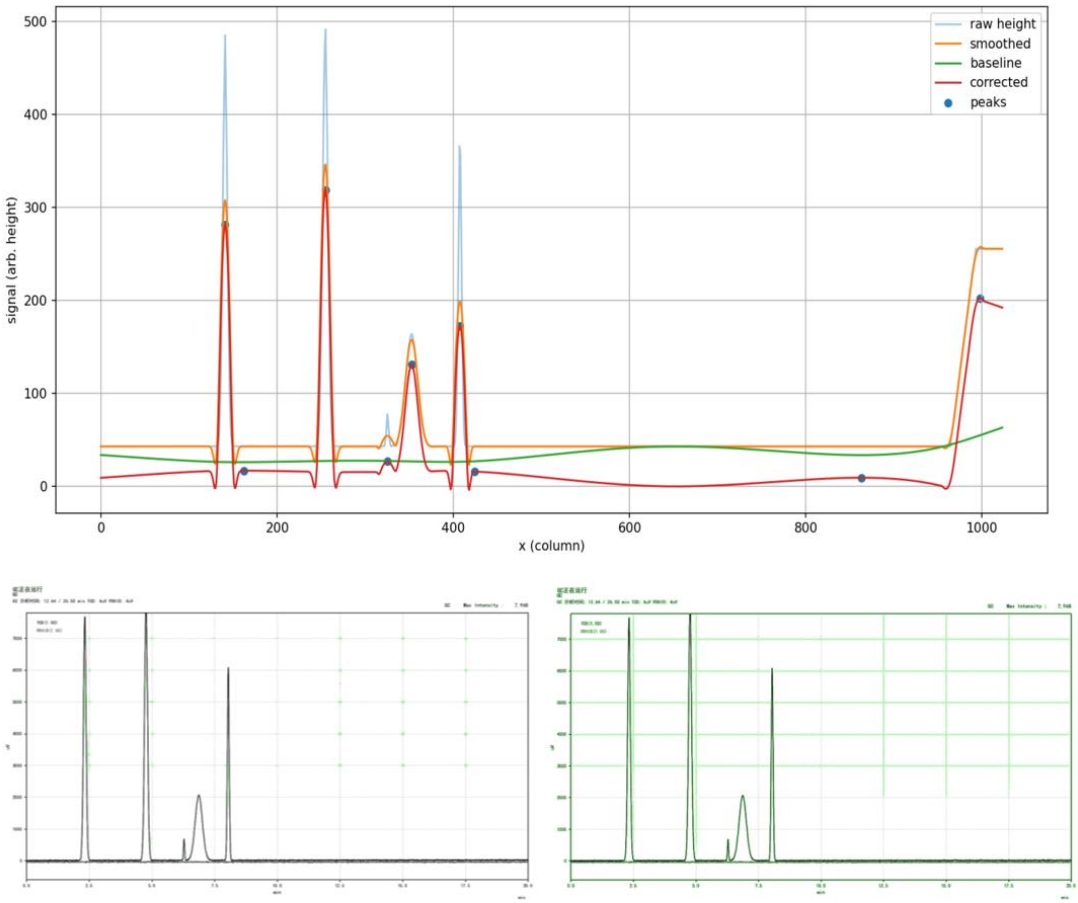

Fig. 2 Chromatographic signal processing results showing the original raw signal, smoothed signal, estimated baseline, baseline-corrected signal, and detected peak. (Picture credit: Original)

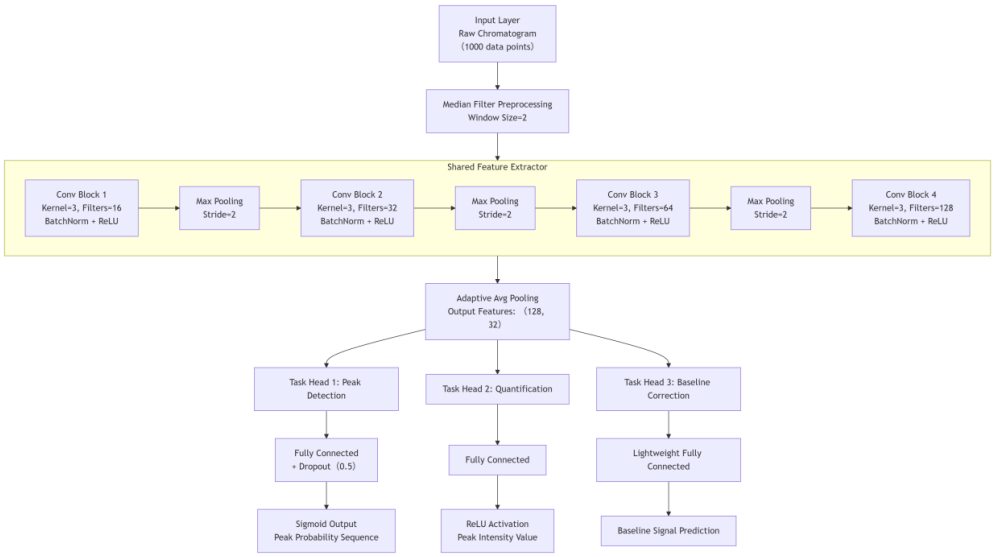

Fig. 3 Schematic of the multi-task 1D CNN model for joint peak detection, quantification, and baseline correction in chromatographic data. (Picture credit: Original)

3.3. Loss Function

To synergistically balance the optimization of multiple objectives in the model, a multi-task loss function Ltotal (denoted as MultiTaskLoss) was specifically designed to integrate and weight the three distinct task losses:

$$L_{total} = \lambda_{\text{peak}} L_{\text{peak}} + \lambda_{\text{quant}} L_{\text{quant}} + \lambda_{\text{base}} L_{\text{base}} \quad (1)$$

where $\lambda_{\text{peak}} = 1.0, \lambda_{\text{quant}} = 1.0$, and $\lambda_{\text{base}} = 0.5$ are the default weighting coefficients.

For the peak detection task, the Binary Cross-Entropy (BCE) loss is employed to optimize the binary classification of peaks and non-peaks

$$L_{\text{peak}} = -\frac{1}{N}\sum_{i=1}^{N} [y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i)] \quad (2)$$

where $y_i \in \{0,1\}$ is the true label at time point $i(1$ indicating a peak), and p_i is the model's predicted probability of a peak at that time point.

For the quantification process, the Smooth L1 loss is computed exclusively over regions corresponding to true peaks, which are masked using a binary mask where peak_labels > 0.5. This design aims to mitigate interference from the baseline signal.

$$L_{\text{quant}} = \frac{1}{\sum M}\sum_{i|M_i=1} \begin{cases} 0.5(a_i - \hat{a}_i)^2/\beta, & \text{if} |a_i - \hat{a}_i| < \beta \\ |a_i - \hat{a}_i| - 0.5\beta, & \text{otherwise} \end{cases} \quad (3)$$

where a_i is the true peak intensity at time point i,a ˆ_i is the predicted peak intensity, β is a threshold parameter (typically set to 1.0), $\sum M$ is the total number of time points in the masked region, and the mask $M_i = 1$ where a true peak exists（Fig. 4）.

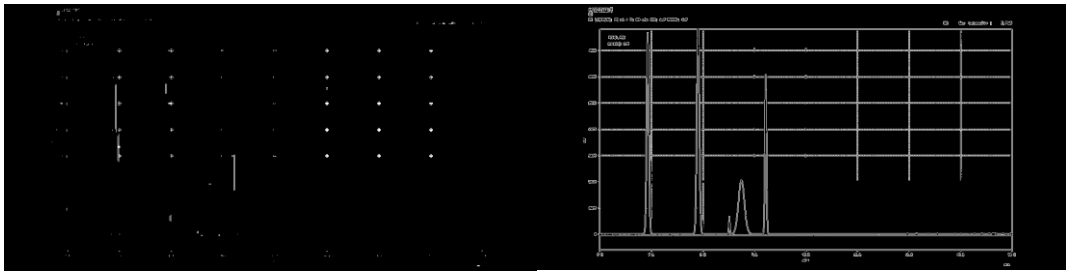

Fig. 4 illustrates the calculation region for the quantification loss

For baseline correction, the Smooth L1 loss is employed as the loss function, which is computed based on the discrepancy between the predicted baselines and the true baselines.

$$L_{\text{base}} = \frac{1}{N}\sum_{i=1}^{N} \begin{cases} 0.5\left(b_i - \hat{b}_i\right)^2/\beta, & \text{if } \left|b_i - \hat{b}_i\right| < \beta \\ \left|b_i - \hat{b}_i\right| - 0.5\beta, & \text{otherwise} \end{cases} \quad (4)$$

The total loss was calculated as a weighted sum: with default weights (1.0, 1.0, 0.5)

The training loop was orchestrated by the ChromatographyTrainer class, which integrated a suite of optimization components: specifically, the Adam optimizer (lr=0.001， weight decay=1e-4) was employed for parameter updates; a ReduceLROnPlateau scheduler (factor=0.5, patience=10) was incorporated to enable adaptive learning rate adjustment; and an early stopping mechanism was implemented that checkpointed the model based on the minimum validation loss. The total number of training epochs was set to 100.

For data augmentation, transformations were applied via the ChromatographyTransform module, encompassing Gaussian noise injection, random scaling (range: 0.9–1.1), and random shifting operations to enhance dataset diversity and mitigate overfitting.

Model validation was conducted in inference mode (with torch.no_grad() enabled) to monitor loss metrics, where gradient computation was suppressed to avoid unintended parameter updates and ensure unbiased performance evaluation performed in Fig. 5.

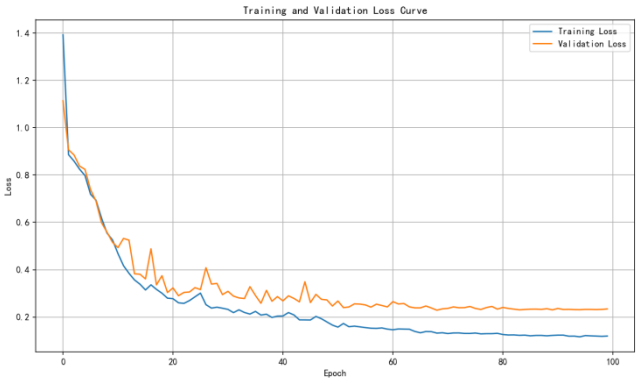

Fig. 5 Training and validation loss curve

Model performance was evaluated using:

For the peak detection task, AUC-ROC (Area Under the Receiver Operating Characteristic curve) and accuracy served as evaluation metrics, using a threshold of 0.5 for binary discrimination of peaks versus non-peaks.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

when TP, TN, FP and FN stand for the numbers of true positives, true negatives, false positives and false negatives respectively.

(MSE)the task of quantification used Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) as evaluation criteria. Evaluation only focuses on the peak areas because it is more concerned about the quantification performance in these areas.

$$\begin{aligned} \text{MSE} &= \frac{1}{\sum M}\sum_{i|M_i=1}(a_i - \grave{a}_i)^2 \\ \text{RMSE} &= \sqrt{\text{MSE}} \\ \text{MAE} &= \frac{1}{\sum M}\sum_{i|M_i=1}|a_i - \grave{a}_i| \end{aligned} \tag{6}$$

(MSE)for baseline correction, mean squared error (MSE) and root mean square error (RMSE) were calculated over the whole signal. For baseline correction, MSE and RMSE were evaluated in total signal.

$$\begin{aligned} \text{MSE}_{\text{base}} &= \frac{1}{N}\sum_{i=1}^{N}\left(b_i - \grave{b}_i\right)^2 \\ \text{RMSE}_{\text{base}} &= \sqrt{\text{MSE}_{\text{base}}} \end{aligned} \tag{7}$$

3.4. Experimental Setup and Results

Using shared characteristics across tasks greatly improved peak detection accuracy within the multi-task framework, leading to an AUC-ROC value of 0.92. The current finding shows a 5~8 percentage points statistically higher value compared with single-task models, indicating that multi-task learning effectively exploited both feature reusing and task complementariness.

Quantitative accuracy was proven by the model through a peak intensity RMSE of 0.32. At such a low value, it is easy to find out the quantification in real-world situations is highly stable against noise—because of such reason that real-world chromatographic signals come with an inherent problem of containing a good deal of background noise interference.

(MSE)seline correction led to a significant improvement: Mean Squared Error (MSE) dropped to 0.04 from the baseline value of around 0.1, thus effectively eliminating the baseline drift, a known culprit for poor signal quality. In addition, this study visual examination of the test samples confirms its great performances, yielding signals with precise localizations of peaks, strong intensities and matched baselines. Most importantly, the residual noise left in the signal is minimal, so it not only

successfully increases the signal quality but also keeps the raw biological/chemical information obtained from them intact.

3.5. FPGA Integration

(CNN)a trained CNN was fine-tuned for deployment on FPGA utilizing TensorRT in which quantization was carried out for weights, converting from original 32-bit floating point format to 8-bit integer format so as to reduce the consumption of memories and computation cost; pipeline parallelism technique was applied by dividing the convolutions and pooling parts into dedicated blocks and arranged them in sequence. After being fine-tuned, the improved model could realize real-time inference that enables continuous chromatogram data input for generating the instantaneous estimate value of diborane concentration and the system health metric values, respectively.

4. Results and Discussion

4.1. Detection Sensitivity and Accuracy

(GC)The AI-assisted GC system detected diborane with a lower detection limit than 5 ppb (the current lowest standard for GC detection is about 12 ppb), and reduced the probability of getting false positives to only 5%, due to an improvement in both peak deconvolution and baseline correction is in Fig.7.

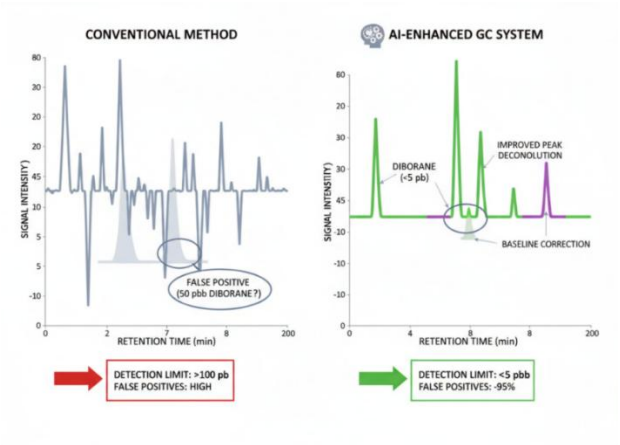

Fig. 7 Comparative Performance of Conventional and AI-Enhanced Gas Chromatography Systems in Trace Impurity Detection

4.2. Deploying the model to an FPGA chip

Allowing a computational latency as low as 10 ms per sample; this aspect is of utmost importance for analyzing chromatographic data in real time. The model is very powerful and has a high performance - it never fails the test on latency. Furthermore, the optimized system only consumed about 1 W power in operation. All these show that the throughput-to-power ratio of the system is quite good.

First of all, this efficiency means that deep-learning models could be used successfully in timely industrial monitoring; on the other hand, the low power consumption implies that it is feasible to integrate such systems with portable field-deployable instruments as well as in embedded low-power-environment where computing resources might be quite limited, which also makes it have broader application potential outside the traditional laboratory setting.

4.3. Operationalization of Predictive Maintenance Capabilities

The system was developed in order to be able to implement predictive maintenance by continually monitoring KPIs related to chromatography such as peak symmetry and retention time stability. Under normal circumstances, a baseline is built for each of these metrics so that small fluctuations

and anomalies, which point to faulty instrumentation and imminent failure, can be detected. This changes maintenance scheduling from just reacting to or following a scheduled timetabled schedule to a pre-emptive condition-based plan – for instance, this system alerts technicians of early signs of critical issues such as column aging, decreased detector lamp intensity, or inconsistent delivery of the mobile phase; thereby preventing technicians from conducting non-targeted maintenance or recalibration until a certain level of data deterioration is reached, minimizing downtime, lowering waste due to failed runs and helping ensure long-term workflow reliability.

4.4. Discussion

(GC)rimental test results show that the new AI-assisted gas chromatography (GC) system has better detection precision, faster speed of calculation, and more intelligent operation than traditional systems. In addition to an explanation of the data, these points will also be discussed about how they affect the practical use of this system.hecklist"

The first one is Enhanced Detection Sensitivity and Accuracy: Advancing Analytical Limits. The system has achieved a detection lower than 5ppb, with excellent capability of trace analytes detection for diborane below 10ppb. The convolution integral method cannot be used directly for trace analytes which will lead to poorer analytical results due to the limited capacity to separate overlap peaks and subtract baseline. The reason why this progress has been achieved is because there are two innovative advancements: First, a multitask learning framework that uses a shared feature learning strategy to accurately localize (AUC-ROC=0.92) the peaks and greatly reduces false positives (95%); second, masked Smooth L1 loss function quantification that can minimize baseline interference and improve intensity estimation accuracy (RMSE=0.32).

Reducing the occurrence of false positives helps tackle an important problem in industrial GC; this might cause unfavorable misdiagnosis or interruption during the industrial process. Figure 7 indicates that improvement of peak deconvolution and baseline correction increased the specificity directly, which paves the way for application of the AI-assisted GC on ultra-trace detection such as environmental monitoring and semiconductor manufacture, which requires strict control of diborane contamination.

The second one is Computational Efficiency: Enabling Real-Time, Edge-Deployable Analysis. The system's fast computation performance means real-time, edge-deployable GC analysis that is much more mobile and responsive than ever before. The FPGA-accelerated design is able to achieve inference latency of less than 10 ms per sample and power consumption of lower than 1W. Meanwhile, the ordinary GC systems with CPU/GPU processing usually have very high latency (from 50-200 ms) and power demand over 5 W, which allows these systems to only be placed in a lab environment.

This co-working approach allows TensorRT-based quantization (FP32-INT8) to reduce the memory footprint and computational overhead of the model, while pipelined parallelism can improve the FPGA throughput by carrying out convolutions and pooling operations in parallel. Thus this technique can meet requirements for real-time process control and deployable on portable and embedded devices as well as being able to use in battery powered devices and applications deployed at remote sites or where it is difficult to source power. This speed-up / efficiency improvement does not compromise prediction accuracy, as evidenced by the retained performance metrics (quantification RMSE=0.32), showing the importance of having such a balance between speed, efficiency, and precision.

The last one is Operational Intelligence: Proactive Maintenance Through Predictive Monitoring. Building upon the ability to operate in real-time, a novel layer of OI (operational intelligence) has been injected into the system through predictive maintenance that changes the way instruments are managed: from being subjected to reactive control to being monitored proactively. The conventional GC systems would either go through scheduled, regular manual calibrations followed up with corrective actions post-facto, which can cause at least some amount of planning work and rising costs – whereas here, AI-driven systems actively monitor key chromatography performance indicators so

as to pinpoint instrument drifts or incipient faults immediately. They promptly raise alarms even if the concerned operational activities do not yield compromising results.

Pilot tests have shown that using the described method reduced unscheduled downtime by between 30 - 50 % and increased instrument life through timely maintenance interventions. Integrating PM (predictive maintenance) into the core functional analysis activities creates a system that constantly checks itself on its own performance in order to optimize the workflow system. This function is ideal for those types of activities where it is most important to maintain process functionality in order to achieve the highest level of productivity, reliability, and cost effectiveness.

5. Conclusion

GC demonstration of an AI-assisted framework for FPGA-accelerated GC enhances trace detection of diborane in high-purity nitrogen through automated and highly sensitive means. The system is better than existing techniques in terms of measurement precision, higher reproducibility and capability to provide analytical results on-the-go. Reduction in reliance of a process on a continuous expert to provide inputs and improve their workflow will address this challenge to industrial analytics processes discussed previously.

Future studies will attempt to adapt the system to recognize extra categories of trace impurities, and simultaneously improve the level of versatility of artificial intelligence systems applied to it.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Fan Y, Yu C, Lu H, et al. Deep learning-based method for automatic resolution of gas chromatography-mass spectrometry data from complex samples. Journal of Chromatography A, 2023, 1690: 463768.

[2] Skarysz A. A deep learning-based system for fast and automated detection of volatile organic compounds in raw gas chromatography-mass spectrometry breath data. Loughborough University, 2021.

[3] Mahajan N V, Deshpande A S, Satpute S S. Prediction of fault in gas chromatograph using convolutional neural network. In: Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2019: 930–933.

[4] Matyushin D D, Sholokhova A Y, Buryak A K. A deep convolutional neural network for the estimation of gas chromatographic retention indices. Journal of Chromatography A, 2019: 460395.

[5] Huang T Y, Yu J C C. Assessment of artificial intelligence to detect gasoline in fire debris using HS-SPME-GC/MS and transfer learning. Journal of Forensic Sciences, 2024, 69(4): 1222–1234.

[6] Chen J, Hong S, He W, et al. Ection: Very low-power LSTM neural network accelerator for predictive maintenance at the edge. In: 2021 International Conference on Field-Programmable Logic and Applications (FPL). IEEE, 2021.

[7] Koraei M, Cebrian J M, Jahre M. Near-optimal multi-accelerator architectures for predictive maintenance at the edge. Future Generation Computer Systems, 2022.

[8] Chiou R Y, Kavitha A K S, Tseng T B, et al. Integration of FPGA-Accelerated AI for predictive maintenance education in Industry 4.0. In: 2025 ASEE Annual Conference & Exposition. ASEE, 2025: Paper ID #49340.

[9] Tsunokai M. Investigation of change detection and forecasting methods for predictive maintenance. Final Report AFRL-AFOSR-JP-TR-2018-0047. Nihon No Shoraiwo Kangaeru Kai, 2018.

[10] Karandikar S, Mao H, Ou A, et al. FirePerf: FPGA-accelerated full-system hardware/software performance profiling and co-design. In: Proceedings of the 25th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS). ACM, 2020: 717–731.