chemosensors

IMPACT FACTOR 3.7

CITESCORE 5.0

Article

A Deep Learning Approach to Investigating Clandestine Laboratories Using a GC-QEPAS Sensor

Giorgio Felizzato, Nicola Liberatore, Sandro Mengali, Roberto Viola, Vittorio Moriggia and Francesco Saverio Romolo

Special Issue

Chemical Sensing and Analytical Methods for Forensic Applications

Edited by

Prof. Dr. Jorn Yu

MDPI

https://doi.org/10.3390/chemosensors12080152

chemosensors

MDPI

Article

A Deep Learning Approach to Investigating Clandestine Laboratories Using a GC-QEPAS Sensor

Giorgio Felizzato [1,*], Nicola Liberatore [2], Sandro Mengali [2], Roberto Viola [2], Vittorio Moriggia [3] and Francesco Saverio Romolo [1,*]

[1] Department of Law, University of Bergamo, Via Moroni 255, 24127 Bergamo, Italy
[2] Consorzio CREO, 67100 L'Aquila, Italy; nicola.liberatore@consorziocreo.it (N.L.); sandro.mengali@consorziocreo.it (S.M.); roberto.viola@consorziocreo.it (R.V.)
[3] Department of Economics, University of Bergamo, Via dei Caniana 2, 24127 Bergamo, Italy; vittorio.moriggia@unibg.it
* Correspondence: giorgio.felizzato@unibg.it (G.F.); francescosaverio.romolo@unibg.it (F.S.R.)

Abstract: Illicit drug production in clandestine laboratories involves the use of large quantities of different chemicals that can be obtained for legitimate purposes. The identification of these chemicals, including reagents, catalyzers and solvents, is crucial for forensic investigations. From a legal point of view, a drug precursor is a material that is specific and critical to the production of a finished chemical and that constitutes a significant portion of the final molecular structure of the drug. In this study, a gas chromatography quartz-enhanced photoacoustic spectroscopy (GC-QEPAS) sensor—in conjunction with a deep learning model—was evaluated for its effectiveness in the detection and identification of interesting compounds for the production of amphetamine, methamphetamine, 3,4-methylenedioxymethamphetamine (MDMA), phenylcyclohexyl piperidine (PCP), and cocaine. The GC-QEPAS sensor includes a gas sampler, a fast GC for separation, and a QEPAS detector, which excites molecules exiting the GC column using a quantum cascade laser to provide the infra-red (IR) spectrum. The on-site capability of the GC-QEPAS system offers significant advantages over the current instruments employed in this field, including rapid analysis, which is crucial in field operations. This allows law enforcement to quickly identify specimens of interest on site. The system's performance was validated by taking into account the limit of detection, repeatability, and within-laboratory reproducibility. The results showed excellent repeatability and reproducibility for both the GC and QEPAS modules. The deep learning model, a multilayer perceptron neural network, was trained using IR spectra and retention times, achieving very high classification accuracy in the testing conditions. This study demonstrated the efficacy of the GC-QEPAS sensor combined with a deep learning model for the reliable identification of drug precursors, providing a robust tool for law enforcement during criminal investigations in clandestine laboratories.

Keywords: machine learning; drug precursor; analytical instrument; gas chromatography; quartz-enhanced photoacoustic spectroscopy

Citation: Felizzato, G.; Liberatore, N.; Mengali, S.; Viola, R.; Moriggia, V.; Romolo, F.S. A Deep Learning Approach to Investigating Clandestine Laboratories Using a GC-QEPAS Sensor. *Chemosensors* **2024**, *12*, 152. https://doi.org/10.3390/chemosensors12080152

Received: 1 July 2024
Revised: 28 July 2024
Accepted: 3 August 2024
Published: 5 August 2024

1. Introduction

A major source of illicit drugs today is the clandestine laboratory, where substances such as 3,4-Methylenedioxymethamphetamine (MDMA), phenylcyclohexyl piperidine (PCP), and methamphetamine are manufactured. Many processes can be used to manufacture controlled substances, depending on the desired end product and the starting materials. Most of the chemicals used in the manufacture of drugs have legitimate uses and can be obtained without restrictions through chemical suppliers, supermarkets, and shops. The key to criminal investigations in clandestine laboratories is recognizing combinations of chemicals that can occur in controlled substances. The three main types of chemicals used in the manufacturing of controlled substances are reagents, catalyzers, and solvents [1,2].

Reagents are chemicals used to induce a reaction, typically involving one of the precursors. They contribute only a small portion of the end product. Solvents are liquids used to solubilize reagents and are used as carriers during a reaction. They are also used to purify end products [3]. The United Nations Office on Drugs and Crime refers to precursors as materials specific and critical to the production of a finished chemical, representing a significant portion of the final molecular structure of the drug [4]. According to the EU precursors COUNCIL REGULATION (EC) No 111/2005, precursors are "scheduled" (controlled) and are divided into several categories: category 1 covers the most sensitive substances from which illicit drugs can be produced most easily; category 2 covers less sensitive substances; and category 3 covers bulk chemicals that can have different types of uses in the manufacturing process (e.g., feedstock, solvents, impurity removers, etc.) [5].

Clandestine laboratory investigation is one of the most dangerous tasks undertaken by law enforcement due to the presence of hazardous chemical compounds [2]. A "bomb factory" cannot be immediately distinguished from a clandestine laboratory preparing drugs of abuse, despite the presence in the scientific literature of analytical approaches to spot bomb factories [6–10].

Many portable devices have been developed to mitigate the danger of crime scene investigations in clandestine laboratories. These instruments not only assist in uncovering illicit drug laboratories and chemical waste dumps but also enable the detection and identification of volatile organic compounds (VOCs), which safeguards the health of operators [11]. Deena et al. [12] developed a colorimetric sensor that uses machine learning to detect drugs and their precursors. Montiel et al. [13] demonstrated the effectiveness of an electrochemical method based on the electrochemical fingerprint of amphetamine-type stimulant precursors by employing a derivatization approach that enabled the electrooxidation of benzyl methyl ketone (BMK). Wen et al. [14] developed a portable embedded drug precursor gas detection device based on a cataluminescence-based sensor array that involved a chemometric approach to analyzing drug precursor patterns. Colling et al. [15] investigated the efficiency and effectiveness of a microfluidic gas-to-liquid interface for the extraction of target amphetamines and their precursors from air samples.

This study shows that a gas chromatography quartz-enhanced photoacoustic spectroscopy (GC-QEPAS) sensor combined with deep learning is effective in the detection and identification of common drug precursors, reagents, and solvents used in illicit drug production. Although this is not the first time that a QEPAS-based sensor has been used to analyze precursors of drugs of abuse, this is the first time that a QEPAS approach based on artificial neural networks (ANNs) specifically developed for on-site use of the sensor during criminal investigations in clandestine laboratories has been used [16].

ANNs are at the core of deep learning, as they are versatile, powerful, and scalable, making them well suited for carrying out large and highly complex tasks [17]. ANNs are computing systems whose structure is inspired by biological neural networks. Generally, they consist of many simple processors linked by weighted connections. By analogy, the processing nodes can be called neurons. The power of the system emerges from the appropriate combination of many units [18].

Typically, an ANN consists of three main components: the input layer, which receives data; the hidden layers, which are responsible for recognizing patterns associated with the analyzed process or system; and the output layer, which comprises neurons that generate the final outputs derived from prior layer processing. Among the ANN architectures, the multilayer perceptron (MLP) is the most prevalent. It has a multilayer feedforward design in which information flows unidirectionally from the input layers to the output layers (Figure 1). MLP was chosen for this study because its classification efficiency has been proven to be effective and reliable in many fields [19–22].

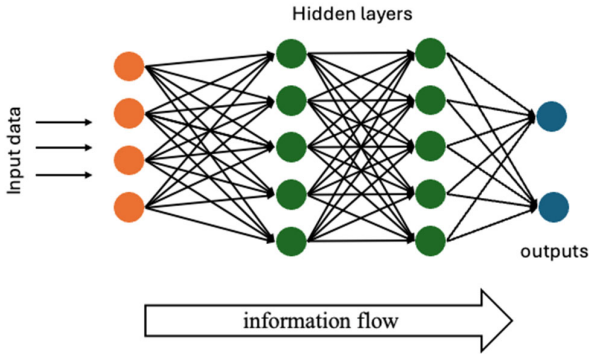

Figure 1. Multilayer perceptron scheme.

Indeed, MLP is expected to provide better results than traditional methods that rely on comparing retention times and infrared (IR) spectra with databases. The conventional approach, which involves comparing IR spectra to reference spectra in databases, typically relies on similarity metrics such as the correlation coefficient and often struggles to deal with the variability introduced by different experimental conditions. By contrast, as a neural network, MLP can automatically learn and extract features from the training experimental dataset. This capability allows the MLP to build a robust model, taking into account the variability in spectra due to different sample matrices, sample concentrations, noise, and other experimental conditions. By learning from the data itself, the MLP method builds a more adaptable and reliable model. The robustness of the MLP model is particularly advantageous in real-world scenarios, such as monitoring clandestine laboratories in which experimental conditions can vary significantly. Traditional database comparison methods may struggle in these scenarios because they cannot handle the complexity and variability of real data [23].

2. Materials and Methods

In this study, six chemicals used for the production of illicit drugs were analyzed: safrole (category 1-EU regulation, CAS 94-59-7), piperidine (category 1-EU regulation, CAS 110-89-4), BMK (category 1-EU regulation, CAS 103-79-7), methyl ethyl ketone (MEK; category 3-EU regulation, CAS 78-93-3), acetone (category 3-EU regulation, CAS 67-64-1), toluene (category 3-EU regulation, CAS 108-88-3), and benzaldehyde (CAS 100-52-7). All the target analytes were purchased from Monforte Lab Suppliers (Grassobbio, Italy).

Benzaldehyde and MEK are involved in the synthesis of methamphetamine through the Baeyer–Villinger reaction [24]. BMK, or 1-phenyl-2-propanone, is the precursor most commonly used for amphetamine and methamphetamine synthesis [25]. Four principal precursors—safrole, isosafrole, piperonal, and 3,4-methylenedioxyphenyl-2-propanone (PMK)—are typically used in the illicit manufacture of MDMA and related drugs. Moreover, safrole is the key starting material, as the other three precursors can be synthesized from it [26]. Piperidine is a common component in the production of the hallucinogen PCP ("angel dust") [27]. Acetone and toluene are extensively employed as solvents in various illicit drug production processes, including cocaine purification. MEK is also used as a solvent in cocaine production [28].

To achieve a precise concentration of the target analytes in air, a 60 L glass chamber was connected to the sensor sampler via a plastic tube. The samples were introduced into the chamber by puncturing a porous septum located on the tap of the box using a 10 µL syringe (Eppendorf, Hamburg, Germany). Below the septum, a heated crucible was installed to facilitate sample evaporation within seconds. To prevent sample decomposition, the temperature of the hot crucible was carefully controlled, maintaining a sufficiently high level to enhance analyte vaporization without causing thermal degradation. The average temperature used was around 40 °C. Additionally, the results obtained with the hot crucible were compared to those from analyses conducted without it, focusing on both retention

time and the shape of the IR curve, to ensure no sample degradation occurred. Pearson's correlation coefficient was used to compare the IR curves. As a result, a well-defined concentration of analytes was achieved within the chamber.

Analyses were carried out by introducing more than one compound into the chamber in order to reproduce a more realistic scenario in which more target compounds are expected to be present.

Multivariate data analysis was performed using Python code within a Jupyter Notebook environment running on an Apple Mac Mini equipped with an 8-core CPU, a 10-core GPU, and 8 GB of unified memory. The following packages were used in the modules: NumPy [29], Pandas [30], Matplot Library [31], Plotly [32], and Scikit-Learn [33].

2.1. Instrument

This article examines the effectiveness of a GC-QEPAS sensor (Consorzio CREO, Italy) in detecting and identifying illicit drug precursors [16]. The sensor comprises three main components:

- a gas sampler and pre-concentrator designed to deal with large air volumes (a compact purge and trap device based on commercial sorbent tubes from Markes International Ltd. that can sample approximately 1 L of air and transfer pre-concentrated vapors to the FAST-GC separation module in less than 3 min).
- a fast GC (CNR-IMM, Bologna, Italy) [34] consisting of one micro-electro-mechanical system (MEMS) for preconcentration and injection and one MEMS GC column for separation integrated on a silicon micro-machined chip.
- a QEPAS detector that measures the photoacoustic spectra of the analytes eluted by the fast GC and that incorporates a quantum cascade laser source (MiniQCL, Block Engineering, Southborough, MA, USA) capable of continuously scanning the thermal IR spectrum within wavelengths ranging from 7.4 μm to 10.7 μm for spectroscopic analysis.

Additionally, the sensor is equipped with a mini-PC controller to manage all the sensing chains automatically and run spectral analysis algorithms for identification. The GC-QEPAS sensor was used with the settings shown in Table 1. The sensor parameters used in this article were already optimized in [16].

Table 1. GC-QEPAS sensor settings.

Pre-Concentrator Stage	
Sampling time	10–60 s
Sampling flow	800–1000 mL/min
Release temperature	200–290 °C
GC settings	
Initial temperature	60–80 °C
Hold time	5–30 s
Ramp rate	120–140 °C/min
Max temperature	240–270 °C
Hold time	1–2 min
QEPAS settings	
QTF cell temperature	80–120 °C
Laser wavelength scan	8–10 μm
Laser modulation	32,760–32,768 Hz
Laser pulse width	200 ns

2.2. Validation

The analytical method underwent a validation process following the European Network of Forensic Science Institute guidelines for the analysis of illicit drugs and the Commission Implementing Regulation (EU) 2021/808 based on a validation plan for qualitative methods [35,36]. For the qualitative method developed in this study, the following validation parameters were taken into account: limit of detection (LOD), repeatability, and within-laboratory reproducibility.

LOD values were experimentally estimated by considering three times the signal-to-noise ratio of the recorded chromatographs.

Repeatability was calculated both for the chromatographic outputs (retention time), using the standard deviation and the coefficient of variation (CV%) of the retention time, and the IR spectrum, considering the position of the maximum value in the spectrum. Repeatability was calculated based on at least six replicates of the analysis of each target analyte performed on the same day and under the same experimental conditions.

Within-laboratory reproducibility was determined for both the chromatographic results (retention time, t_R) and the IR spectra. The standard deviation and CV% of the retention times were evaluated to assess the chromatographic outputs. Regarding the IR spectra, the position of the maximum of the spectra was evaluated using the standard deviation and CV%. Reproducibility was calculated based on a minimum of 18 analysis replicates for each target analyte. Replicates were spread across at least three different days, with six replicates carried out per day by varying operators, reagent and solvent batches, room temperature, and ambient moisture levels.

3. Results and Discussion

The validation parameters were computed separately for both the GC and QEPAS modules to ensure the reliability of the data obtained from each component of the sensor system. This separation ensures that each module's performance can be individually assessed and optimized. Table 2 shows the LOD and repeatability of the GC outputs.

Table 2. Limit of detection and repeatability of the GC outputs.

Analyte	Average t_R (s)	SD t_R (s)—Repeatability	CV% t_R—Repeatability	LOD (ppm)
Acetone	65.36	0.97	1.5	0.96
Methyl ethyl ketone	75.40	1.3	1.7	2.4
Toluene	83.15	0.81	0.98	0.066
Piperidine	84.54	1.0	1.2	0.68
Benzaldehyde	132.1	1.4	1.1	0.0026
Benzyl methyl ketone	179.7	2.2	1.2	0.12
Safrole	182.6	1.2	0.65	0.24

Benzaldehyde had the lowest LOD among the target analytes (estimated at 0.0026 ppm). Repeatability and within-laboratory reproducibility of the GC module were assessed based on the CV% of retention times. The CV% of retention times was below 2% for all analytes on the same day (repeatability). Within-laboratory reproducibility showed slightly higher CV% values, with acetone nearing 3% and toluene at 2.24%, but remained acceptable for reliable analytical performance. Notably, for all other analytes, a CV% lower than 2% was obtained for this parameter. The results indicate that variations in experimental conditions had a minimal impact on retention times (Table 3).

Table 3. Within-laboratory reproducibility of the GC outputs.

Target Analyte	Mean Retention Time (s)	SD (s)	CV%
Acetone	62.96	1.9	3.1
Methyl ethyl ketone	74.68	1.4	1.9
Toluene	83.69	1.9	2.2
Piperidine	85.32	1.5	1.8
Benzaldehyde	134.3	2.1	1.6
Benzyl methyl ketone	179.3	1.9	1.0
Safrole	184.9	2.1	1.1

To validate the QEPAS module, the position of the maximum of the spectra was taken into account. The CV% for repeatability was consistently below 0.5% for all analytes. Notably, for three analytes (acetone, BMK, and safrole), there were no variations among the IR spectra (recorded at 0.00%, rounded to two decimal places) (Table 4). Similar results were obtained for within-laboratory reproducibility. The CV% for all analytes remained below 0.5% (Table 5).

Table 4. Repeatability of QEPAS spectra.

Target Analyte	Average Peak Position (Maximum of the Spectrum) (μm)	SD (mm)	CV%
Acetone	8.18	0.00	0.0
Methyl ethyl ketone	8.48	0.10	0.12
Toluene	9.66	0.048	0.49
Piperidine	8.93	0.0045	0.050
Benzaldehyde	8.26	0.0067	0.081
Benzyl methyl ketone	8.14	0.00	0.00
Safrole	9.48	0.00	0.00

Table 5. Within-laboratory reproducibility of QEPAS spectra.

Target Analyte	Average Peak Position (Maximum of the Spectrum) (μm)	SD (mm)	CV%
Acetone	8.18	0.0043	0.053
Methyl ethyl ketone	8.49	0.030	0.35
Toluene	9.63	0.042	0.44
Piperidine	8.92	0.0071	0.079
Benzaldehyde	8.26	0.0066	0.080
Benzyl methyl ketone	8.14	0.0065	0.080
Safrole	9.48	0.0078	0.082

The comprehensive validation procedure underscores the excellent repeatability and reproducibility of the analytical results achieved by both the GC and QEPAS systems. The retention times of various targets have been shown to be repeatable and reproducible. Variations in operators and experimental conditions had minimal impacts on the final results.

A similar validation protocol was applied to the QEPAS module to ensure the reliability of its outputs. The CV% for each target analyte confirms the good results of both repeatability and within-laboratory reproducibility. Therefore, the analytical method is

effective for its intended purpose and can be confidently used to investigate clandestine laboratories that produce illicit drugs.

Classification via Deep Learning

The IR spectra obtained from the QEPAS module underwent standard normal variate pre-processing, which involved centering and scaling each spectrum based on its standard deviation to mitigate scattering effects [37]. The MLP ANN was then trained using three different inputs. First, only the pre-processed IR spectra were used. Second, the model was trained using both outputs from the sensor: retention times and IR spectra. Finally, principal component analysis (PCA) was used to reduce the dimensionality of the training set, which comprised retention times and IR spectra, resulting in new variables called principal components (PCs). These components were also used to train the MLP model.

The MLPs in this study offer an accurate and robust classification approach by considering the inherent variability and complex patterns within the data. Moreover, MLP was shown to be an effective choice for investigating clandestine laboratories where the experimental conditions vary significantly, thereby making them particularly suitable for real forensic scenarios.

The MLP models were tuned for the three input approaches using two hidden layers, each containing one hundred neurons. Moreover, two different solvers were tested: lbfgs, which is typically used for small datasets, and adam, which is useful for large datasets (i.e., with thousands of training samples or more) [33]. Six different model configurations were evaluated. As a result, six different model configurations were evaluated, considering the following combinations of inputs and solvers: (1) IR spectra using adam solver; (2) IR spectra using lbfgs solver; (3) IR spectra coupled with the corresponding retention times using adam solver; (4) R spectra coupled with the corresponding retention times using lbfgs solver; (5) PCs based on IR spectra and retention times using adam solver; and (6) PCs based on IR spectra and retention times using lbfgs solver.

The parameters for the estimator, using both the lbfgs and adam solvers, were obtained with the "model.get_params" function. The details are shown in Table 6.

Table 6. Parameters for MLP model using lbfgs and adam solver.

Parameters	Value
Activation	Relu
Batch_size	Auto
Beta_1	0.9
Beta_2	0.999
Epsilon	1×10^8
Hidden_layers_size	100
Learning_rate	Constant
Max_fun	15,000
Max_iter	200
Power_t	0.5
Random_state	None
Verbose	False

The models' accuracy was assessed to identify the most effective strategy for classifying the GC-QEPAS data. Leave-p-out cross-validation (LPO; class sklearn.model_selection. LeavePOut(p)) was selected as the primary method for evaluating the models' performance. LPO randomly splits the original dataset into a training set, which is used to set up the model, and an evaluation set, which is used to assess its performance [38]. In this study,

70% of the dataset was used as the training set, and the remaining 30% was used as the evaluation set. Table 7 shows the accuracy of the models.

Table 7. Model accuracy with three different types of input.

Input Data	**Model Accuracy**	
	Solver: adam	Solver: lbfgs
IR spectra	100%	100%
IR spectra and t_R	100%	100%
PCs based on IR spectra and t_R	100%	93%

The models trained using only IR spectra and those trained using a combination of IR spectra and retention times demonstrated identical performance, achieving a final accuracy of 100% with no misclassifications. This consistent accuracy across different inputs indicates the robustness of the MLP models. Overfitting was considered unlikely, as the models were validated using an independent evaluation set comprising a different range of samples from those used in the training set.

Furthermore, the model trained with PCA scores achieved 100% accuracy when using the adam solver but exhibited lower accuracy when using the lbfgs solver, where BMK was completely misclassified as safrole. This significant misclassification suggests that the use of PCA scores with the lbfgs solver may not be suitable for real-world applications. Therefore, the approach using PCA scores with the lbfgs solver is not recommended for practical scenarios. The high accuracy achieved with the adam solver should be considered reliable only when the concentration of the target compounds is well above the LoD. Near the LoD, the accuracy of the deep learning models is expected to decrease. The lowest concentration of target analytes used for training the deep learning models in this study was greater than one unit of measurement above the LoD. This approach was deemed reliable because the concentration of chemicals in a clandestine laboratory is expected to be higher than LoD because criminals use pure compounds to obtain high yields of products.

Additionally, the time taken to train the models and make predictions using the evaluation set was calculated to evaluate the efficiency of each solver and input configuration. The fastest solver for training the model was lbfgs when using the principal components derived from the IR spectra and retention time as input. However, for prediction speed, this model configuration was the second fastest. For prediction, the adam solver demonstrated the quickest performance when using the IR spectra as input data. This highlights the adam solver's effectiveness in scenarios where rapid predictions are crucial, despite its slightly slower training time compared to the lbfgs solver. Table 8 provides a detailed comparison of the training and prediction times for different input data and solver combinations.

Table 8. Time taken to train the models and make predictions.

Training Time		
Input data	Solver: adam	Solver: lbfgs
IR spectra	0.386 s	0.445 s
IR spectra and t_R	0.0567 s	0.0331 s
PCs based on IR spectra and t_R	0.0600 s	0.0207 s
Prediction Time		
Input data	Solver: adam	Solver: lbfgs
IR spectra	0.0004 s	0.0248 s
IR spectra and t_R	0.0009 s	0.011 s
PCs based on IR spectra and t_R	0.032 s	0.0007 s

In summary, while the MLP models trained with either IR spectra alone or in combination with retention times showed excellent accuracy and robustness, the performance of the PCA-based approach varied significantly depending on which solver was used. These findings highlight the importance of selecting appropriate pre-processing techniques and solvers to ensure the accurate classification of GC-QEPAS data.

4. Conclusions

This study demonstrated the effectiveness of the GC-QEPAS sensor combined with the MLP deep learning model in the on-site detection and identification of chemicals associated with illicit drug production. MLP based on IR spectra and spectra plus retention times was shown to be a better classification method than using PCA scores to train the models. MLP can automatically learn and extract relevant features from the experimental data, dealing with the inherent variability and complexity of real-world scenarios. This capability ensures high accuracy and robustness across a wide range of experimental conditions. The MLP can handle different sample matrices, concentrations, and noise levels, making it more reliable for use in environments such as clandestine laboratories, where many conditions are unpredictable. The integration of the MLP with the GC-QEPAS sensor not only enhanced its detection capabilities but also ensured that the system remained effective and accurate across a wide range of scenarios. This robustness demonstrated the suitability of the MLP approach for on-site forensic applications, highlighting its potential for broader implementation in detecting and identifying chemical substances under various conditions.

The GC-QEPAS sensor ensured accurate and efficient on-site analysis of various drug precursors, reagents, and solvents involved in the production of street samples of amphetamine, methamphetamine, MDMA, PCP, and cocaine. The validation results highlight the sensor's high sensitivity, repeatability, and reproducibility, showing a low LOD and robust performance across different experimental conditions. The MLP models trained using the IR spectra alone or in combination with retention times proved to be highly effective in classifying the chemicals, achieving nearly 100% accuracy in the experimental conditions adopted to simulate the forensic scenario of a clandestine drug laboratory.

The comprehensive validation and successful application of the GC-QEPAS sensor highlight its potential as a powerful tool that law enforcement agencies can use in crime scene examinations of clandestine drug laboratories. By facilitating the detection and identification of critical chemicals on site, this approach can enhance the capabilities of investigators and provide them with timely information that can increase the safety and efficacy of police investigations of illegal drug-related cases.

Author Contributions: Conceptualization, G.F., N.L., S.M., R.V., V.M. and F.S.R.; methodology, G.F., N.L., S.M., R.V., V.M. and F.S.R.; software, G.F., N.L., S.M., R.V., V.M. and F.S.R.; validation, G.F., N.L., S.M., R.V., V.M. and F.S.R.; formal analysis, G.F. and F.S.R.; investigation, G.F., N.L., S.M., R.V., V.M. and F.S.R.; resources, G.F., N.L., S.M., R.V., V.M. and F.S.R.; data curation, G.F., V.M. and F.S.R.; writing—original draft preparation, G.F., N.L., S.M., R.V., V.M. and F.S.R.; writing—review and editing, G.F., N.L., S.M., R.V., V.M. and F.S.R.; visualization, G.F., N.L., S.M., R.V., V.M. and F.S.R.; supervision, N.L., S.M., R.V. and F.S.R.; project administration, S.M. and F.S.R.; funding acquisition, N.L., S.M., R.V. and F.S.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by RISEN project, HORIZON2020 grant number 883116.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study will be made available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tilstone, W.J.; Hastrup, M.L.; Hald, C. *Fisher Techniques of Crime Scene Investigation First International Edition*; CRC Press: Boca Raton, FL, USA, 2019.
2. Christian, D.R., Jr. *Forensic Investigation of Clandestine Laboratories*; CRC Press: Boca Raton, FL, USA, 2022.
3. United Nations Office on Drugs and Crime. *The International Drug Control Conventions: Tables of the United Nations Convention against Illicit Traffic in Narcotics Drugs and Psychotropic Substances of 1988*; United Nations: New York, NY, USA, 2019.
4. United Nations Office on Drugs and Crime. *World Drug Report 2013*; United Nations Office on Drugs and Crime: Vienna, Austria, 2013; pp. 55–71.
5. *COUNCIL REGULATION (EC) No 111/2005 of 22 December Laying Down Rules for the Monitoring of Trade between Precursors 2004 the Community and Third Countries in Drug Precursors*; European Union: Brussel, Belgium, 2023.
6. Önnerud, H.; Wallin, S.; Östmark, H.; Menning, D.; Ek, S.; Ellis, H.; Kölhed, M. Localisation of threat substances in urban society-LOTUS: A viable tool for finding illegal bomb factories in cities. In *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense X*; SPIE: Bellingham, WA, USA, 2011; Volume 8019.
7. Fiorani, L.; Puiu, A.; Rosa, O.; Palucci, A. Lidar/DIAL detection of bomb factories. In *Electro-Optical Remote Sensing, Photonic Technologies, and Applications VII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing*; SPIE: Bellingham, WA, USA, 2013; Volume 8897.
8. Romolo, F.S.; Connell, S.; Ferrari, C.; Suarez, G.; Sauvain, J.-J.; Hopf, N.B. Locating bomb factories by detecting hydrogen peroxide. *Talanta* **2016**, *160*, 15–20. [CrossRef]
9. Desmet, C.; Degiuli, A.; Ferrari, C.; Romolo, F.S.; Blum, L.; Marquette, C. Electrochemical sensor for explosives precursors' detection in water. *Challenges* **2017**, *8*, 10. [CrossRef]
10. Ferrari, C.; Ulrici, A.; Romolo, F.S. Expert system for bomb factory detection by networks of advance sensors. *Challenges* **2017**, *8*, 1. [CrossRef]
11. Man, G.; Stroeber, B.; Walus, K. *Detecting and Identifying Clandestine Drug Laboratories: Sensing Technology Assessment*; Canadian Police Research Centre/Centre Canadien de Recherches Policières: Regina, SK, Canada, 2008; Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32005R0111 (accessed on 30 June 2024).
12. Francis, D.P.; Laustsen, M.; Dossi, E.; Treiberg, T.; Hardy, I.; Shiv, S.H.; Hansen, B.S.; Mogensen, J.; Jakobsen, M.H.; Alstrøm, T.S. Machine learning methods for the detection of explosives, drugs and precursor chemicals gathered using a colorimetric sniffer sensor. *Anal. Methods* **2023**, *15*, 2343–2354. [CrossRef]
13. Montiel, N.F.; Parrilla, M.; Sleegers, N.; Van Durme, F.; van Nuijs, A.L.; De Wael, K. Electrochemical sensing of amphetamine-type stimulants (pre)-precursors to fight against the illicit production of synthetic drugs. *Electrochim. Acta* **2022**, *436*, 141446. [CrossRef]
14. Sha, W.; Cui, D.; Li, B.; Wang, Q.; Wang, Y.; Zheng, C. A portable embedded drug precursor gas detection and identification device based on cataluminescence-based sensor array. *Sens. Actuators B Chem.* **2014**, *200*, 132–139. [CrossRef]
15. Collins, M.; Gel, M.; Lennard, C.; Spikmans, V.; Forbes, S.; Anderson, A. Application of a microfluidic gas-to-liquid interface for extraction of target amphetamines and precursors from air samples. *Micromachines* **2020**, *11*, 315. [CrossRef] [PubMed]
16. Liberatore, N.; Viola, R.; Mengali, S.; Masini, L.; Zardi, F.; Elmi, I.; Zampolli, S. Compact GC-QEPAS for on-site analysis of chemical threats. *Sensors* **2023**, *23*, 270. [CrossRef]
17. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2022.
18. Dongare, A.D.; Kharde, R.R.; Kachare, A.D. Introduction to artificial neural network. *Int. J. Eng. Innov. Technol. (IJEIT)* **2012**, *2*, 189–194.
19. Heidari, E.; Sobati, M.A.; Movahedirad, S. Accurate prediction of nanofluid viscosity using a multilayer perceptron artificial neural network (MLP-ANN). *Chemom. Intell. Lab. Syst.* **2016**, *155*, 73–85. [CrossRef]
20. Balabin, R.M.; Safieva, R.Z. Gasoline classification by source and type based on near infrared (NIR) spectroscopy data. *Fuel* **2008**, *87*, 1096–1101. [CrossRef]
21. Naresh, M.; Nagaraju, V.S.; Kollem, S.; Kumar, J.; Peddakrishna, S. Non-invasive glucose prediction and classification using NIR technology with machine learning. *Heliyon* **2024**, *10*, e28720. [CrossRef] [PubMed]
22. Aizaldeen Abdullah, A.; Aldhahab, A.; Al Abboodi, H.M. Review of eye diseases detection and classification using deep learning techniques. In *BIO Web of Conferences 97*; EDP Sciences: Les Ulis, France, 2024; p. 00012.
23. Qi, Y.; Hu, D.; Jiang, Y.; Wu, Z.; Zheng, M.; Chen, E.X.; Liang, Y.; Sadi, M.A.; Zhang, K.; Chen, Y.P. Recent progresses in machine learning assisted Raman spectroscopy. *Adv. Opt. Mater.* **2023**, *11*, 2203104. [CrossRef]
24. Doughty, D.; Painter, B.; Pigou, P.E.; Johnston, M.R. The synthesis and investigation of impurities found in clandestine laboratories: Baeyer–Villiger route part I; synthesis of P2P from benzaldehyde and methyl ethyl ketone. *Forensic Sci. Int.* **2016**, *263*, 55–66. [CrossRef] [PubMed]
25. EMCDDA and Europol, EU Drug Market: Amphetamine-in Depth Analysis, October 2023. Available online: https://www.euda.europa.eu/publications/eu-drug-markets/amphetamine_en (accessed on 30 June 2024).
26. United Nations Office on Drugs and Crime. *Ecstasy and Amphetamines Global Survey 2003*; United Nations Office on Drugs and Crime: Vienna, Austria, 2003.
27. European Monitoring Centre for Drugs and Drug Addition (EMCDDA). *Drug Precursor Developments in the European Union*; EMCDDA Papers; Publications Office of the European Union: Luxembourg, 2019.

28. Cunningham, J.K.; Callaghan, R.C.; Liu, L. US federal cocaine essential ('precursor') chemical regulation impacts on US cocaine availability: An intervention time-series analysis with temporal replication. *Addiction* **2015**, *110*, 805–820. [CrossRef] [PubMed]
29. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef]
30. The Pandas Development Team, Pandas-Dev/Pandas: Pandas; version: 2.2.2, Zenodo. 2020. Available online: https://pandas.pydata.org/docs/index.html (accessed on 30 June 2024).
31. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [CrossRef]
32. Plotly Technologies Inc. *Collaborative Data Science*; Plotly Technologies Inc.: Montréal, QC, Canada, 2015.
33. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *22*, 2825–2830.
34. Zampolli, S.; Elmi, I.; Cardinali, G.C.; Masini, L.; Bonafè, F.; Zardi, F. Compact-GC platform: A flexible system integration strategy for a completely microsystems-based gas-chromatograph. *Sens. Actuators B Chem.* **2020**, *305*, 127444. [CrossRef]
35. European Network of Forensic Science Institute. *Best Practice Manual (BPM) for Controlled Drug Analysis*; European Network of Forensic Science Institute: Wiesbaden, Germany, 2020.
36. COMMISSION IMPLEMENTING REGULATION (EU) 2021/808 of 22 March 2021on the performance of analytical methods for residues of pharmacologically active substances used in food-producing animals and on the interpretation of results as well as on the methods to be used for sampling and repealing Decisions 2002/657/EC and 98/179/EC. Available online: https://eur-lex.europa.eu/eli/reg_impl/2021/808/oj (accessed on 30 June 2024).
37. Bi, Y.; Yuan, K.; Xiao, W.; Wu, J.; Shi, C.; Xia, J.; Chu, G.; Zhang, G.; Zhou, G. A local pre-processing method for near-infrared spectra, combined with spectral segmentation and standard normal variate transformation. *Anal. Chim. Acta* **2016**, *909*, 30–40. [CrossRef]
38. Liu, S. Leave-*p*-out cross-validation test for uncertain Verhulst-Pearl model with imprecise observations. *IEEE Access* **2019**, *7*, 131705–131709. [CrossRef]