The Application of Deep Learning in Source Attribution of Ignitable Liquid Residues in Fire Debris for Forensic Fire Investigation

Ting-Yu Huang, PhD[1] and Jorn Yu, PhD[2]

[1] *Department of Fire Science, Central Police University, Taiwan*

[2] *Department of Forensic Science, Sam Houston State University, Huntsville, TX 77340*

ABSTRACT

This study presents a deep learning workflow for source attribution of ignitable liquid (IL) residues in fire debris, using heated headspace solid-phase microextraction combined with GC/MS (HHS-SPME-GC/MS) data. A convolutional neural network (CNN) was trained on pseudo-color chromatographic heatmaps to discriminate among gasoline brands. The model achieved high accuracy on neat and fire-degraded samples, demonstrating robustness, minimal training data requirements, and consistent, automated forensic interpretation.

INTRODUCTION

Source attribution of ignitable liquid (IL) residues in fire debris is a critical component of forensic fire investigation, as it can provide investigative leads by linking residues to potential sources. Gas chromatography–mass spectrometry (GC-MS) is the current analytical standard for identifying and classifying ILs [1]; however, brand-level discrimination remains challenging due to complex chemical profiles, fire-induced degradation, and matrix interferences. Conventional statistical approaches often require extensive feature selection and expert interpretation [2-4].

Recent advances in deep learning (DL) have demonstrated strong capabilities for learning hierarchical patterns from complex datasets, even with limited training data. In particular, convolutional neural networks (CNNs) can extract subtle, reproducible features from chromatographic data when the data are represented as image inputs [5-7].

This study hypothesized that a DL-based workflow trained on neat ILs can learn to focus on consistent chemical features and remain effective at discriminating between IL brands when applied to diluted, fire-degraded, and matrix-interfered samples. The poster presents the experimental design, data preprocessing strategy, model development, and evaluation framework, along with performance metrics from neat gasoline samples and simulated fire debris samples to assess robustness and generalizability.

MATERIALS & METHODS (1/2)

Table 1. Gasoline database.

Dataset	Neat gasoline samples (Band C and K diluted to various levels)	Simulated fire debris samples (Band C and K diluted to various levels and spiked onto 250 mg burned Nylon carpet)
Training	126	N/A
Verification	36	N/A
Assessment	N/A	36
Total	198	

RESULTS & DISCUSSION

Transfer learning

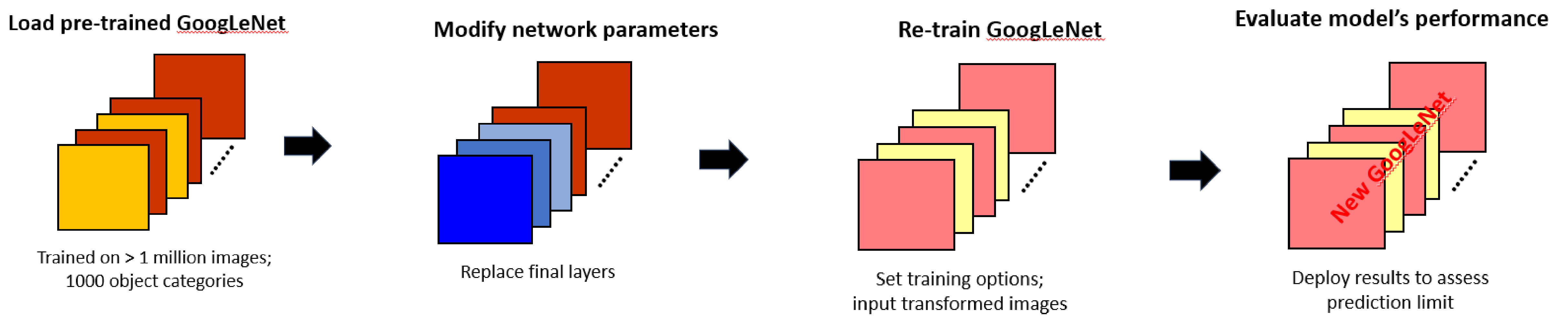

Image transformation

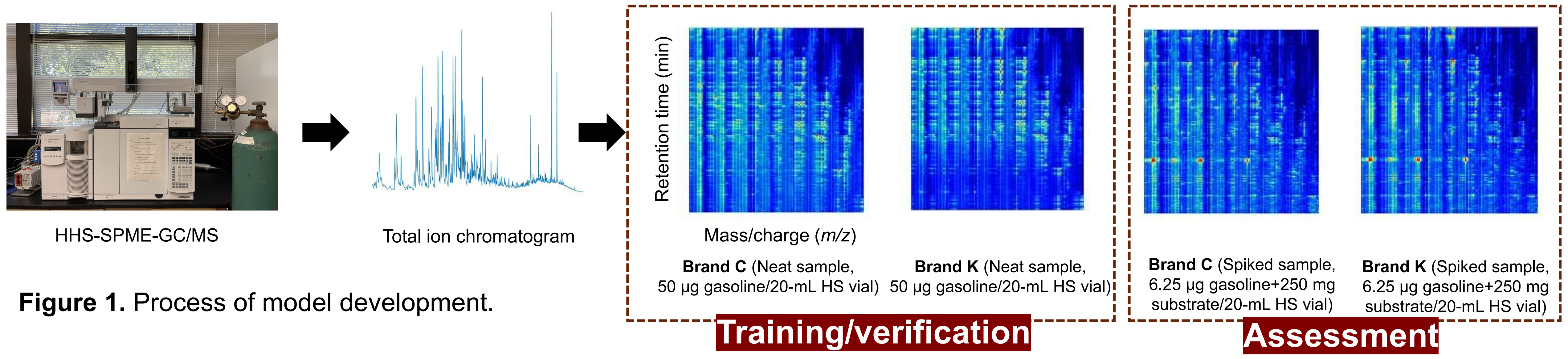

Figure 1. Process of model development.

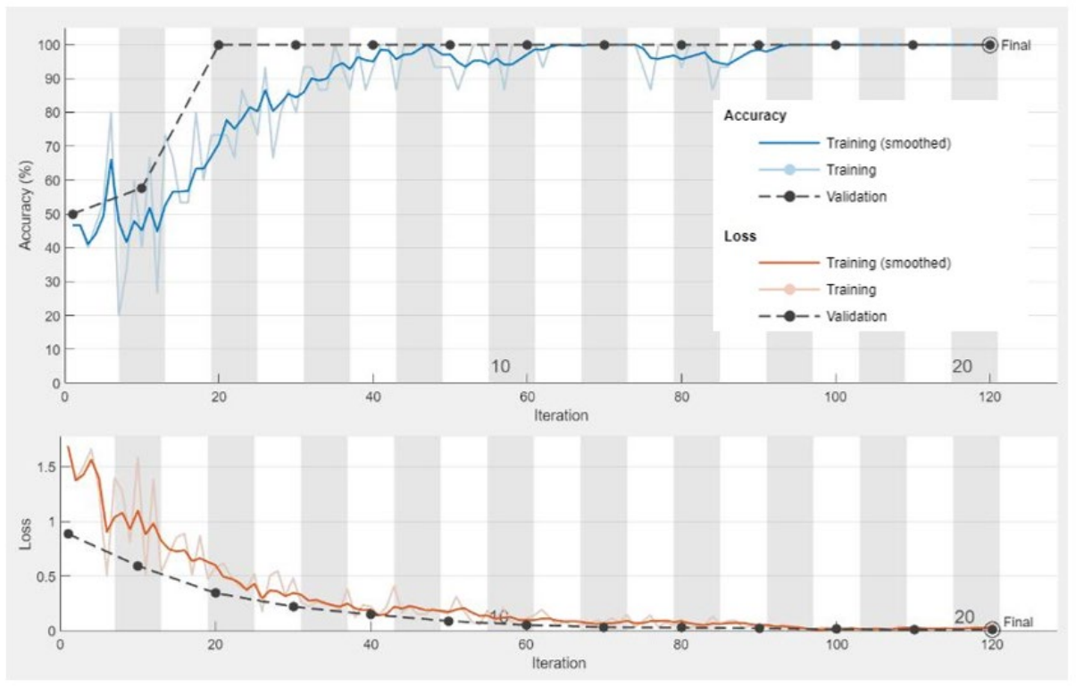

Figure 2. Training progress and validation results.

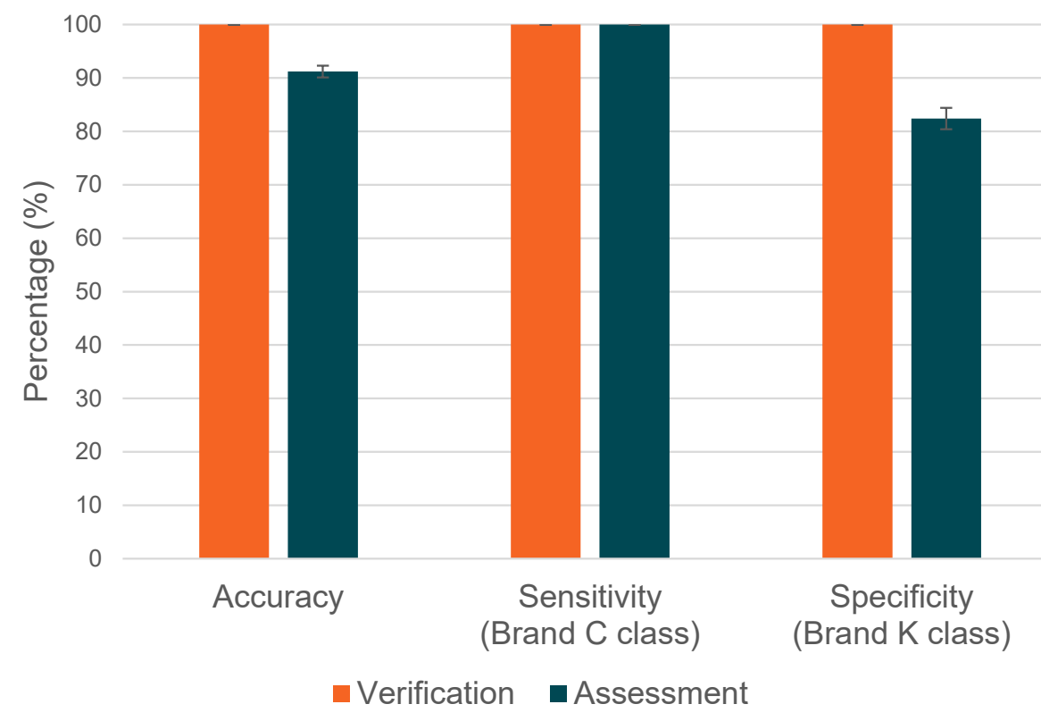

Figure 3. Verification and assessment results.

- LOD (Neat gasoline sample): 0.39 µg per 20-mL HS vial.
- LOD (Simulated fire debris sample): 6.25 µg per 20-mL HS vial.

Table 2. GC/MS settings.

Oven program steps	Condition
GC oven initial temperature	40 °C
Hold time	2 min
Rate #1, Oven temperature #1, Hold time #1	10 °C/min, 150 °C, 0 min
Rate #2, Oven temperature #2, Hold time #2	30 °C/min, 300 °C, 0 min

Table 3. HHS-SPME settings.

HHS-SPME step	Condition
Pre-fiber conditioning temperature	250 °C
Pre-fiber conditioning time	60 s
Pre-incubation time	300 s
Incubation temperature	80 °C
Extraction time	120 s
Desorption time	120 s
Post-fiber conditioning temperature	250 °C
Post-fiber conditioning time	600 s

REFERENCES

[1] American Society for Testing and Materials. Standard Test Method for Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography-Mass Spectrometry. https://store.astm.org/e1618-19.html. Accessed 1 July 2025.

[2] Shadan AF, Juahir H. Identification and prediction of fresh gasoline locations and branding using newly targeted compound chromatograms with chemometrics and machine learning. Unesa Journal of Chemistry. 2023;12(2):73–82. doi:10.26740/ujc.v12n2.p73-82.

[3] Monfreda M, Gregori A. Differentiation of unevaporated gasoline samples according to their brands, by SPME-GC-MS and multivariate statistical analysis. J Forensic Sci. 2011;56(2):372-80. doi: 10.1111/j.1556-4029.2010.01644.x.

[4] Suppajariyawat P, De Andrade AFB, Elie M, Baron M, Gonzalez-Rodriguez J. The use of chemical composition and additives to classify petrol and diesel using gas chromatography–mass spectrometry and chemometric analysis: a UK study. Open Chem. 2019;17(1):183–97. doi: 10.1515/chem-2019-0021.

[5] Venkateswara SM, Padmanabhan J. Deep learning based agricultural pest monitoring and classification. Sci. Rep. 2025;15(1). doi:10.1038/s41598-025-92659-5.

[6] De T Guerra R, Yamaguchi CK, Stefenon SF, Coelho LDS, Mariani VC. Deep learning approach for automatic heartbeat classification. Sensors. 2025;25(5):1400. doi: 10.3390/s25051400.

[7] Islam MSB, Sumon MSI, Majid ME, Kashem SBA, Nashbat M, Ashraf A, et al. ECCDN-Net: A deep learning-based technique for efficient organic and recyclable waste classification. Waste Management. 2024;193:363–75. doi: 10.1016/j.wasman.2024.12.023.

MATERIALS & METHODS

Sample preparation: (1) Neat gasoline samples: Gasoline samples with 87 octane, from two brands (C and K), were collected at gas stations in Houston, Texas. Stock solutions were made by dissolving 20 mg of each brand in 1 mL of methanol. Working solutions were then prepared by serially diluting the stock in methanol to concentrations ranging from 78 to 20,000 µg/mL (N = 9). For analysis, 5 µL of each calibrator sample was transferred to a 20-mL headspace (HS) vial (Supelco Inc.) for HS-SPME-GC/MS analysis; (2) Simulated fire debris samples: Nylon carpet samples were purchased from a home improvement store in Houston, Texas. A 16 cm^2 piece of the carpet was burned with a butane torch (Bernzomatic, Chilton, WI) for 1 min in air. Then, 5 µL of each calibrator sample was added to 250 mg of the burned carpet in a 20-mL HS vial for HS-SPME-GC/MS analysis.

Instrumental analysis: A total of 198 HHS-SPME-GC/MS data were collected. The database comprises a training, verification, and assessment dataset, as shown in Table 1. An Agilent 7890B gas chromatograph coupled with a 5975A mass spectrometer (Agilent Technologies, Santa Clara, CA) was used to perform HHS-SPME-GC/MS analysis of neat gasoline and simulated fire debris samples. The HHS-SPME-GC/MS settings are detailed in Tables 2 and 3.

GC/MS data processing and transfer learning: The GC/MS data were initially converted to NetCDF format using Agilent ChemStation (Agilent Technologies, Inc., California, USA) for further processing in MATLAB (MATLAB 2023b and Bioinformatics Toolbox, MathWorks, Natick, Massachusetts, USA). To assist feature recognition and differentiate between the sample headspace chemical signatures of Brands C and K during transfer learning, retention times from 3.8 to 12.3 minutes and associated *m/z* values from 50 to 156 were selected from the resampled signal to generate heatmaps for each GC/MS data. These 2D GC/MS visualizations were created as pseudo-color heatmaps that illustrate spectral intensities after applying a log transformation within the selected m/z and retention time ranges (Figure 1). Those images were fed into a pre-trained CNN, GoogLeNet. The re-trained GoogLeNet was used to discriminate "Brand C" and "Brand K" samples. The transfer learning settings are shown in Table 4. Transfer learning was also performed in MATLAB.

Table 4. Transfer learning settings.

Analysis type		Condition
Transfer learning	**Mini batch size**	10
	Max epochs	20
	Initial learn rate	1e-4
	Validation frequency	6

CONCLUSIONS

- Pseudo-color heatmaps retained the characteristic features of gasoline chemical profiles across various brands, enabling effective transfer learning with GoogLeNet.
- Retraining GoogLeNet did not need manual feature extraction.
- GoogLeNet performed well with fewer than 150 training samples.

ACKNOWLEDGEMENTS

This work was partly funded by the 2022-2023 Forensic Sciences Foundation (FSF) Lucas Research Grants and the Ministry of Education (MOE). The opinions, findings, and conclusions or recommendations expressed in this manuscript are those of the author(s) and do not necessarily reflect those of the FSF and MOE. Ting-Yu Huang also appreciates the funding support from the National Science and Technology Council (NSTC), Taiwan.