Electronic Supplementary Material (ESI) for Journal of Analytical Atomic Spectrometry.

Supplementary material

Here, the maximum soil samples of 7 heavy metals were selected, and the XRF and visNIR spectra were used for the GAS conversion process. The spectrum is preprocessed and converted by GASF or GADF to complete the drawing of two-dimensional images.

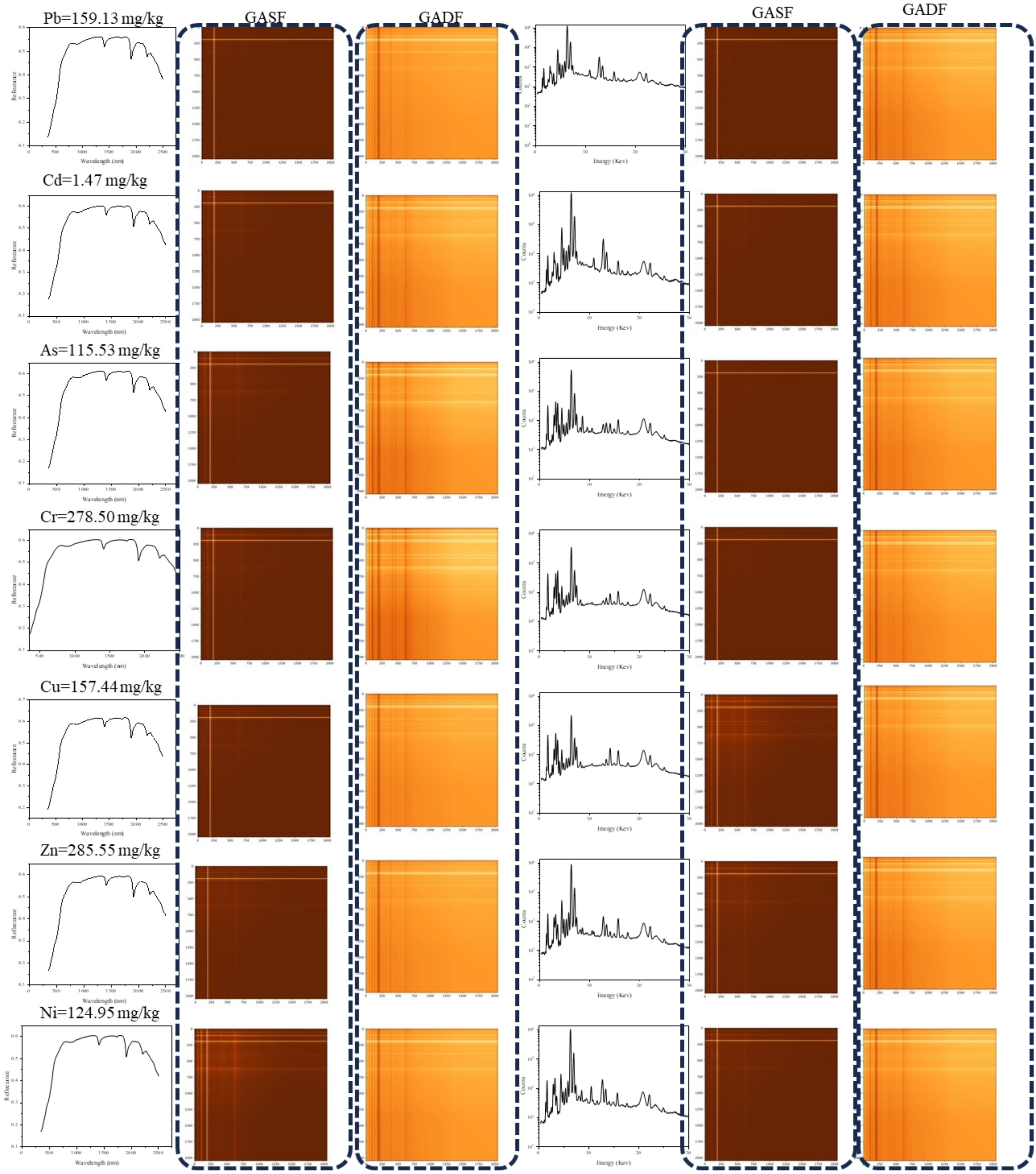

Figure S 1 GAS conversion of visNIR and XRF data from maximum soil heavy metals samples. Seven soil samples were selected, and the corresponding maximum contents of seven heavy metals were Pb=159.13 mg/kg, Cd=1.47 mg/kg, As=115.53 mg/kg, respectively. Cr=278.50mg/kg, Cu=157.44 mg/kg, Zn=285.55mg/kg, Ni=124.95mg/kg. In the visNIR spectrum, the differences are smaller for each sample. There is a large difference between the spectral peaks in XRF spectrum.

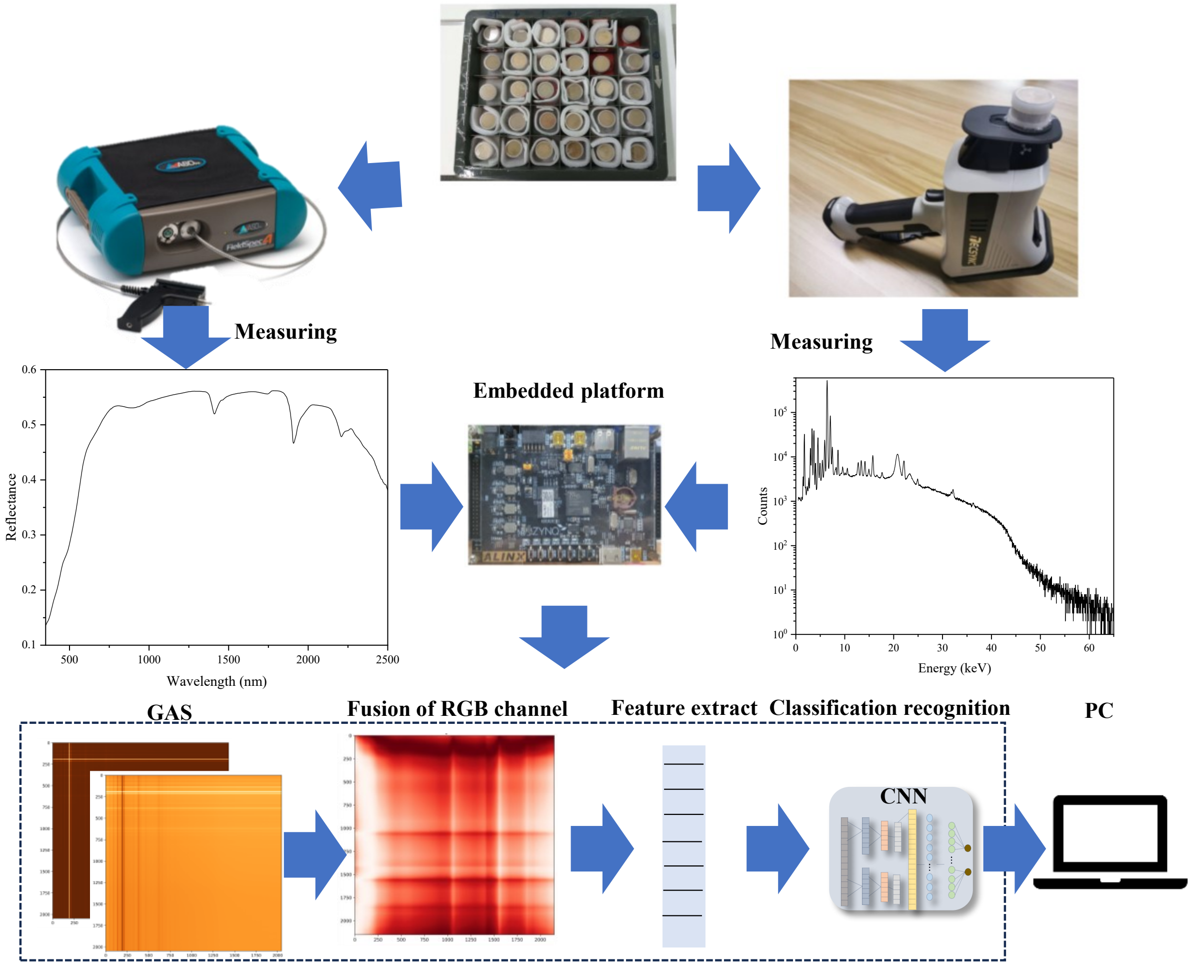

Figure S 2 Measurement and identification of soil by a handheld XRF and visNIR spectrometer.

Soil samples are collected and brought back to the laboratory for necessary pretreatment before being tested using XRF and visNIR spectrometers. After obtaining the spectral data, they are inputted into an embedded platform to undergo a series of processes, including GAS transformation, merging of the RGB channels, PCANet feature extraction, and CNN classification training. The chosen embedded platform is an FPGA development board (Zynq-7000) provided by Shanghai Yiyi Electronic Technology Co., Ltd., China. This development board comprises two ARM Cortex-A9 processor cores which are responsible for running the software of the processor system. It also includes programmable logic, memory, peripheral interfaces, debugging and programming interfaces, as well as power management components. By deploying the pre-trained model onto the embedded platform, the feature variables can be directly recognized. The identification results are then outputted to the computer for further analysis. This approach enables the rapid identification and classification of features in soil samples, providing a convenient and efficient solution for soil heavy metal analysis.

Table S 1 The hyperparameters and the algorithms' training details.

Algorithms	Selected Parameters (Implementation by Sklearn or Pytorch)
1D-CNN	**Input:** 1D data of fusion XRF-visNIR, **Batch size=**64, **Max-epochs=**100, **Optimiser=**Adam, **Learning rate=**1 × 10^{-4}, $\beta_{1=}$ 0.99, $\beta_{2=}$ 0.999
RF	**Input:** 1D data of fusion XRF-visNIR, **max_depth** = 50, **n_estimators** = 50
KNN	**Input:** 1D data of fusion XRF-visNIR, **n_neighbours** = 2, **algorithm**='brute', **weight**='distance'
SVM	**Input:** 1D data of fusion XRF-visNIR **Kernel**='linear', probability = True
XGBoost	**Input:** 1D data of fusion XRF-visNIR, **γ** = 0.1, **max_depth** = 50, **learning_rate** = 0.4
MLP	**Input:** 1D data of fusion XRF-visNIR, **alpha** = 0.5, **hidden_layer_sizes** = 900
GBDT	**Input:** 1D data of fusion XRF-visNIR, **max_depth** = 20, **n_estimators** = 80, **learning_rate** = 0.12

Table S 2 Statistical data table of heavy metal content in each region of Hongfeng Lake

Sampling area	Statistical Parameter	North Lake	Back Lake	South Lake	Center Lake	Independent verification area
Pb	Max	159.13	92.86	124.75	115.94	104.84
	Mean	50.08	42.81	41.19	37.97	44.93
	Min	3.28	1.72	1.40	1.90	6.09
Cd	Max	0.748	0.799	1.470	1.316	0.665
	Mean	0.438	0.411	0.763	0.493	0.397
	Min	0.008	0.164	0.178	0.019	0.059
As	Max	115.53	90.62	78.42	73.44	53.56
	Mean	33.61	36.16	28.58	31.23	23.99
	Min	1.70	3.68	1.83	0.16	2.84
Cr	Max	278.50	191.62	167.21	151.92	198.48
	Mean	123.83	87.47	89.79	91.58	74.19
	Min	5.00	8.49	19.98	35.93	4.14
Cu	Max	97.04	103.27	139.43	157.44	96.77
	Mean	40.53	51.30	56.13	51.36	43.21
	Min	2.91	4.83	1.40	0.53	3.47
Zn	Max	248.58	240.46	246.39	273.49	285.55
	Mean	171.91	172.36	154.41	173.42	173.86
	Min	121.48	57.66	39.96	45.16	91.07
Ni	Max	77.67	78.20	124.94	77.90	104.13
	Mean	58.66	59.05	58.14	57.35	55.95
	Min	41.85	38.28	19.17	41.62	5.39

The number of samples from each area of Hongfeng Lake where heavy metal exceedance surpasses the pollution screening values is shown in Table S 3.

Table S 3 The amount of heavy metals exceeded the pollution screening value

Sampling area	Pb (≥70mg/kg)	Cd (≥0.3mg/kg)	As (≥40mg/kg)	Cr (≥150mg/kg)	Cu (≥50mg/kg)	Zn (≥200mg/kg)	Ni (≥60mg/kg)
North Lake	8	24	8	10	8	7	12
Back Lake	7	20	12	4	16	10	13
South Lake	8	26	9	2	23	9	21
Center Lake	4	29	9	2	19	10	16
Independent verification area	8	27	27	3	12	10	16

The data in the table shows varying levels of heavy metal pollution in different areas of Hongfeng Lake. In North Lake area, the heavy metals are primarily concentrated in Cd and Ni, exceeding the limits. In Back Lake, Cd and Cu are the predominant heavy metals. In South Lake area, Cd and Cu are the main pollutants. In Center Lake, Cd, Cu, and Ni are the major contributors to heavy metal contamination. Even in the independent verification area, which is further away from Hongfeng Lake, heavy metal pollution is still present, with Cd and As being the primary contaminants.