

Identity Verification from Human Scent using Channel Representation of 2D Gas Chromatography-Mass Spectrometry Data

Radim Spetlik¹  Jan Hlavsa¹  Jana Čechová²  Petra Pojmanová² 
Jiří Matas¹  Štěpán Urban² 

¹ Czech Technical University in Prague, Faculty of Electrical Engineering
² University of Chemistry and Technology Prague, Dept. of Analytical Chemistry

Abstract

This study examines the feasibility of employing raw two-dimensional gas chromatography/time-of-flight mass spectrometry (GC×GC–ToF-MS) data for the purpose of human scent identity verification. Unlike techniques that require expert-driven identification of compounds, our framework transforms each GC×GC–ToF-MS sample into a multi-channel image. A comprehensive assessment has been conducted on ten feature-embedding methods, and five spatial-alignment strategies.

The evaluation is performed on a newly assembled dataset of 252 individuals, comprising 2,528 raw samples and aggregating around 7.5 TB of data. On our dataset, the best method reaches $\approx 53\%$ true positive rate at a 5% false positive rate without explicit spatial registration. Although this performance is below that of well-established biometrics (e.g., iris verification), the results demonstrate the feasibility of raw-odor verification for scenarios where direct line-of-sight or cooperation may be limited and suggest that, for our acquisition/drift regime, explicit registration is not required to reach the best performance.

Code and dataset on github.com.

1. Introduction

Identity verification is a central problem in biometrics and a canonical computer vision decision problem, with numerous applications that require accurate and efficient recognition of individuals. Dominant biometric modalities, such as face, iris, and fingerprint recognition, provide robust performance in controlled scenarios. However, they also typically require the subject’s active participation or direct line of sight, and may not always be suitable for applications where direct physical or visual contact is not feasible. Consequently, the search for alternative biometrics that can operate passively — and without explicit subject cooperation — is a compelling research direction.

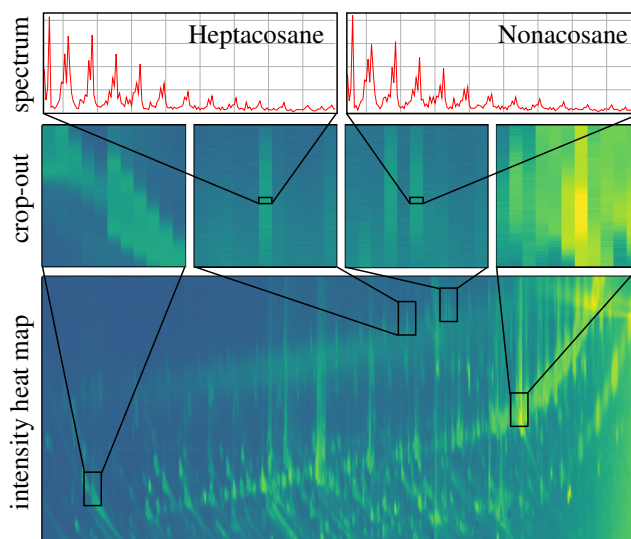


Figure 1. Intensity heat map of the detector response in a 2D Gas Chromatography-Mass Spectrometry measurement, plotted over retention time with t_1 on the x-axis and t_2 on the y-axis (bottom). Mass spectra of Heptacosane and Nonacosane compounds (top). Notably, Heptacosane and Nonacosane spectra are almost identical, with spatial position (retention time) being the primary distinguishing factor. All visualizations are in arbitrary units.

Body odor analysis has long been discussed as a biometric – forensic observations suggest that individuals emit distinctive scent signatures that remain relatively stable over time [20, 22] and can withstand environmental noise from various sources [15, 21, 23, 33], motivating previous attempts at odor-based person recognition using analytical chemistry tools [15, 20] and electronic noses [12, 29]. Existing studies often relied on small, private datasets and expert-driven compound identification [24, 27], limiting scalability and reproducibility. Here, we explore whether modern representation learning can leverage raw outputs from two-dimensional gas chromatography with time-of-

flight mass spectrometry (GC×GC–ToF-MS) to enable large-scale identity verification with minimal expert input.

GC×GC–ToF-MS separates a gas mixture with two capillary columns in series. Each column is a long, narrow tube whose inner wall is coated with a chemical layer called *stationary-phase*. The first column establishes a slow first-dimension retention-time axis t_1 . A periodic modulator (period T_{mod}) slices the t_1 output and injects it into a faster second column, producing a fast axis $t_2 \in [0, T_{\text{mod}})$. The stationary phase chemistry of each column defines its separation pattern; using different coatings across the two columns yields different compound orderings and complementary — often orthogonal — separations. The resulting 2-D chromatogram (t_1, t_2) exhibits peaks that are narrow in t_1 and broad in t_2 (Fig. 1). At each (t_1, t_2) , an electron-impact time-of-flight mass spectrometer (ToF-MS) records an intensity spectrum over mass-to-charge (m/z). The measurement is thus a set of spectra indexed by (t_1, t_2) .

While GC×GC–ToF-MS measurements can capture highly detailed chemical profiles, they historically face two practical barriers: (i) long measurement times (often exceeding 2.5 hours per sample), and (ii) specific sampling procedures (e.g., requiring a subject to touch collecting beads). A recent non-contact sampling device [13] offers a promising direction toward circumventing these limitations: by collecting airborne volatiles passively at a distance, one could potentially avoid visual or physical contact. Such capabilities distinguish scent-based approaches from other biometrics and could eventually enable remote identification where other modalities are not available.

In this work, we represent GC×GC–ToF-MS data as multi-channel images and apply modern feature embedding techniques to the problem of identity verification. In our newly curated dataset of 252 individuals, the best-performing approach yields $\approx 53\%$ true positive rate at a 5% false positive rate. Although this performance is below that of established modalities (e.g., iris recognition), it demonstrates the feasibility of odor-based verification from raw measurements while revealing considerable room for improvement. Our aim is to introduce this problem to the community and encourage further research by publishing our complete dataset (≈ 7.5 TB) and all supporting code.

In summary, our contributions are as follows. We

- (i) propose a new identity verification paradigm that treats raw GC×GC–ToF-MS outputs as multi-channel images, eliminating the need for expert-driven compound identification;
- (ii) introduce a newly assembled, large-scale, open dataset of human-scent GC×GC–ToF-MS measurements (≈ 7.5 TB), covering 252 unique identities;
- (iii) systematically evaluate ten feature-embedding methods, and five data-alignment strategies;
- (iv) demonstrate a prototype identity verification system that

achieves a TPR of $\approx 53\%$ at a 5% FPR, highlighting the feasibility of the approach and the substantial opportunities for further improvement.

2. Related Work

Human body odor has emerged as a distinctive biometric identifier, rooted in the notion that the scent of each individual is unique and relatively stable over time. This idea is supported by long-standing forensic practice: trained dogs can reliably track or identify people by smell, even identical twins who share the same DNA [22]. Scientific analyses have confirmed that individuals emit distinct profiles of volatile organic compounds (VOCs) through sweat and other secretions. For example, Penn et al. conducted a large-scale gas chromatography-mass spectrometry (GC-MS) study of axillary sweat, finding “individually distinct and reproducible GC-MS fingerprints” for each of 197 subjects [20]. In particular, they also observed consistent differences between male and female odor profiles, identifying numerous sex-specific compounds in sweat. This evidence suggests that body odor can serve as a unique “chemical fingerprint” for identity, analogous to a visual fingerprint.

Based on these findings, researchers have developed biometric systems to automate identity verification via scent. Early work applied pattern recognition to personal odor data, demonstrating that people could be identified by their odor with precision of 85% [29]. In that study, an electronic nose (an array of chemical sensors) captured each person’s scent repeatedly (13 subjects over 28 sessions) while learning the odor patterns. The results showed that the odor profile of each subject remained consistent enough for identification, with only an error rate of 15%. This proof-of-concept established odor biometrics as feasible and highlighted its practical appeal: odor-based screening could happen passively and non-invasively. Unlike fingerprint or iris scans that require active cooperation, odor sensors can unobtrusively sample human scent as one walks by. In fact, body odor biometrics have been proposed as a “less aggressive” identification method for settings such as airport security, since the smell of a person could be verified in real time without physical contact or stigmatizing procedures.

2.1. Advancements in Scent Detection Technologies

Recent studies have refined both the sampling sites for human scent and the sensing technologies. One question is which part of the body provides the best odor signature. Breath and underarm sweat are rich in VOCs but pose challenges: breath has high humidity that can interfere with sensors, and armpit odor is often masked by deodorants [2]. In the same study, it is hypothesized that the ear canal could be an effective, relatively uncontaminated source of human odor. Using a portable electronic nose, they collected ear volatiles from volunteers and attempted identification. Re-

markably, their system could identify a person by ear odor with 87% accuracy (10 out of 12 correct) when comparing scent samples taken on the same day. (This greatly exceeds random-chance accuracy of 8.3% for 12 individuals.) However, the study also observed a drop in performance for samples collected on different days – accuracy fell to 22% for day-to-day matching – suggesting that while the odor is unique, its temporal stability and sensor calibration remain challenges for long-term biometric applications. Beyond skin emanations, other work has investigated human breath as an identifier. For example, a study combined a 16-channel chemiresistive sensor array with machine learning to analyze breath samples. With sufficient sensors and data, the authors achieved over 97% identification accuracy on a set of 20 individuals using only their breath odor signatures [12]. These results underscore that, given a robust “digital nose” and algorithms, chemical biometrics can approach the reliability of traditional modalities. In summary, a range of odor-capturing approaches — from GC-MS laboratory analysis to hand-held electronic noses – have proven capable of distinguishing individuals by scent. In [15], human scent residues from fired cartridge cases were reliably characterized using GC×GC–ToF-MS, with chemical profiles remaining stable even after the extreme conditions of a gunshot. In particular, the experiments were carried out on a relatively small dataset involving 20 volunteers, producing statistically robust results (e.g., a Random Forest classifier with an F1 score of 1 and 100% success in Spearman’s correlation similarity comparisons), underscoring the forensic potential of this approach.

2.2. Scent-Based Sex Classification

In addition to identifying who a person is, biological sex can often be inferred from scent alone – a subtopic of great interest in biometrics and forensic science. Early investigations mainly used one-dimensional gas chromatography mass spectrometry (1D GC-MS), using targeted approaches to pinpoint predefined marker compounds indicative of sex, ethnicity, or disease states [3, 5, 9, 14, 20, 28, 31]. The reported accuracies ranged from around 70% to 96.67%, but these findings were based on private datasets and were therefore difficult to replicate [5, 9, 20]. More recently, two-dimensional gas chromatography-time-of-flight mass spectrometry (GC×GC–ToF-MS) has been explored, promising superior compound separation albeit at the cost of more complex data handling. In the first such study, 504 samples collected from 20 women and 20 men were analyzed with a targeted approach, yielding up to 90% accuracy using an SVM classifier [24, 27]. However, the inclusion of samples from the same individuals in both training and test sets limits the generalizability of these results to previously unseen subjects.

2.3. Image-Based Representations

Recent work shows that treating GC×GC–ToF-MS outputs as 2D images rather than isolated peaks can substantially improve performance [1, 8]. These image-based approaches capture subtle chemical differences without manual compound identification, achieving near 100% sensitivity and high specificity in detecting adulterated vs. authentic samples. They also hold promise for identity verification, as the full 2D fingerprint can reveal robust class signatures. Further reviews confirm the suitability of machine learning algorithms for these high-dimensional datasets, with neural networks effectively “smelling” raw chromatograms [4].

2.4. Discussion

Despite these promising results, previous research suffers from several constraints that limit broader adoption and reproducibility. First, almost all studies relied on private or relatively small datasets (often fewer than a few dozen subjects), complicating efforts to validate and compare different methods in larger or more diverse populations [15, 20, 24, 27, 29]. Second, the vast majority of chemical analyses followed a pipeline intensive with human labor, which involved the manual identification and quantification of compounds in 1D or 2D GC-MS experiments [3, 5, 6, 14, 28, 31]. Such manual workflows are prone to operator bias, require specialized chemistry expertise, and often employ targeted search for predefined compounds rather than unbiased, high-throughput discovery. Finally, because of their proprietary nature and specialized hardware, published studies typically do not release raw chemical data in a way that allows the broader scientific community to replicate or extend the findings.

In contrast, we advocate for a fully automated, end-to-end methodology for large-scale odor-based recognition. Rather than relying on manual compound extraction or proprietary pipelines, we employ neural networks to learn robust features, enabling efficient and scalable inference with minimal human intervention. Following a recently published computer vision approach to sex classification [11], we will make our dataset publicly available.

3. Methods

To assess GC×GC–ToF-MS measurements for identity verification, we use a standard verification pipeline. For each input sample, a channel-encoded GC×GC–ToF-MS scan, we compute its embedding. We then compare this embedding to all other embeddings in the dataset. Ideally, embeddings from the same identity form tight clusters, while embeddings from different identities stay well separated.

We extract embeddings from GC×GC–ToF-MS measurements with four approaches: (i) a published baseline method [6] developed by analytical chemists, (ii) hand-

engineered descriptors HOG [7], LBP [18], and SIFT [17], (iii) five off-the-shelf convolutional neural network (CNN) backbones followed by a learned projection layer, and (iv) a custom lightweight CNN that we train from scratch.

Following the sex-classification pipeline introduced in [11], we evaluate five spatial-alignment strategies. We evaluate the four schemes proposed in the original work plus an unaligned variant.

Let $\Omega := \{1, \dots, H\} \times \{1, \dots, W\}$ be the pixel grid and $\mathbf{u} = (u_1, u_2) \in \Omega$. A GC \times GC–ToF–MS measurement is $\mathcal{S} := \{s(\mathbf{u})\}_{\mathbf{u} \in \Omega}$, where the local spectrum at \mathbf{u} is

$$s(\mathbf{u}) = \{(m_{n,\mathbf{u}}, I_{n,\mathbf{u}})\}_{n=1}^{N(\mathbf{u})},$$

with $m_{n,\mathbf{u}}$ the mass-to-charge ratio (m/z) and $I_{n,\mathbf{u}}$ its intensity.

Our Human Scent Dataset includes high- and low-resolution scans. High-resolution produces an irregular, per-pixel set with variable length $N(\mathbf{u})$. Low-resolution uses a fixed m/z grid $\{m_n\}_{n=1}^M$ ($M \approx 800$), yielding

$$s(\mathbf{u}) = \{(m_n, I_{n,\mathbf{u}})\}_{n=1}^M, \quad \text{with } m_n \text{ independent of } \mathbf{u}.$$

3.1. Canonical Frame Registration

Instrument factors, such as temperature changes or hardware drift, can shift retention times across a GC \times GC–ToF–MS measurement, leading to misaligned compounds and unreliable comparisons [11]. To assess alignment explicitly, we evaluate both aligned and unaligned regimes. For the aligned baseline, we reimplemented the canonical-frame registration exactly as in [11], aligning all samples so that the same compounds appear at the same retention times across measurements. For the unaligned setting, we process the raw data directly without any retention-time correction. The registration pipeline is fully automated and requires no manual tuning.

Let $\mathcal{R}^{\text{ref}} = \{c_1, c_2, \dots\}$ denote the set of reference compounds. For each $c \in \mathcal{R}^{\text{ref}}$, let $\mathbf{u}_c \in \mathbb{R}^2$ be the reference location and s_c^{ref} the corresponding mass spectrum. Following [11], we evaluate two subsets reported to occur in all samples of that dataset: $\mathcal{R}^{(22)}$ and $\mathcal{R}^{(24)}$ (22- and 24-compound sets; full lists in Supplementary Sec. S1.2.1). The registration procedure has four steps and is performed for each measurement separately.

1. **Peak Detection** The peak location $\hat{\mathbf{u}}_c^{(S)}$ of every compound c is estimated using one of the compound peak detectors from [11] (for a detailed description of compound peak detectors, see Supplementary Sec. S1.2).
2. **Delaunay triangulation** The Delaunay triangulation is constructed over the estimated locations $\{\hat{\mathbf{u}}_c^{(S)}\}$.
3. **Coordinate transformation** New coordinates for enclosed points are obtained via barycentric interpolation, mapping the estimated compound location $\hat{\mathbf{u}}_c^{(S)}$ to the reference location \mathbf{u}_c .

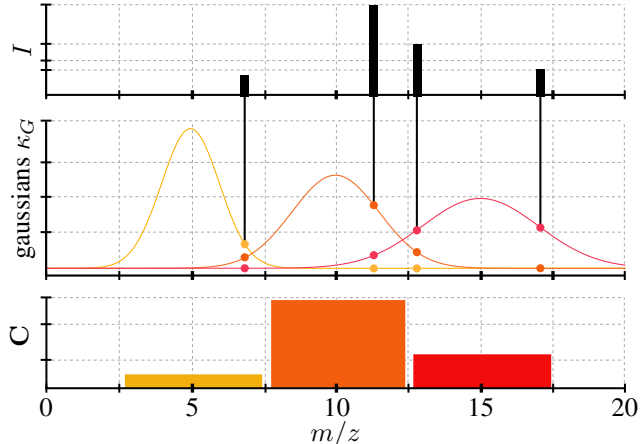


Figure 2. Channel representation of a toy time-of-flight mass spectrometry measurement for $C = 3$ (see Sec. 3.2). The high-resolution spectrum, comprising m/z values and intensities I , is transformed into a compact channel representation \mathbf{C} by applying three basis functions (here Gaussians) with learnable parameters θ and scale factors α . All values are in arbitrary units.

4. **Piecewise-linear interpolation** The original measurement is interpolated onto the transformed mesh to produce a canonical grid.

For a detailed description of the registration process introduced in [11], see Supplementary Sec. S1.3.

3.2. Channel Representation of Time-of-Flight Mass Spectrometry Measurements

We compress the raw mass spectra $s(\mathbf{u})$ with a bank of C overlapping, learnable kernels

$$\phi_k(m_{n,\mathbf{u}}) = \alpha_k \kappa(m_{n,\mathbf{u}}; \theta_k), \quad k = 1, \dots, C, \quad (1)$$

where $\alpha_k > 0$ is a learnable amplitude and $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ is a parametric basis function.

We experiment with the Gaussian kernel function

$$\kappa_G(m_{n,\mathbf{u}}; \theta_k) = \frac{1}{\sqrt{2\pi} \sigma_k} \exp\left(-\frac{(m_{n,\mathbf{u}} - \mu_k)^2}{2\sigma_k^2}\right), \quad (2)$$

where $\theta_k := (\mu_k, \sigma_k)$, with learnable width $\sigma_k > 0$. Fig. 2 illustrates the channel representation with κ_G .

We also experimented with triangular kernels κ_Δ (see Supplementary Sec. S1.4 for a detailed description), which performed worse than Gaussians.

Input Measurement Encoding. We convert \mathcal{S} to a channel-encoded image $\mathbf{C} \in \mathbb{R}^{C \times H \times W}$ by applying the trained kernel bank to each spectrum:

$$\mathbf{C}_{k,\mathbf{u}} = \sum_{n=1}^{N(\mathbf{u})} \phi_k(m_{n,\mathbf{u}}) I_{n,\mathbf{u}}.$$

3.3. Feature Extraction Approaches

In this section, we describe methods for embedding extraction from GC×GC–ToF-MS measurements.

3.3.1. Compound-Centric Baseline

We reimplement the compound-centric baseline [6]. Let $\mathcal{R}^q = \{c_1, \dots, c_Q\}$ be the set of compounds with estimated quantities, $Q := |\mathcal{R}^q|$. A sample \mathcal{S} is represented by $\mathbf{x} \in \mathbb{R}^Q$, where k -th element of \mathbf{x} is the measured quantity of c_k . We apply ℓ_1 normalization: $\mathbf{x} \leftarrow \mathbf{x}/\|\mathbf{x}\|_1$.

Feature selection. For each compound c , we compare absolute quantity differences in within-identity pairs to those in across-identity pairs using a nonparametric two-sample rank test, yielding a p -value p_c . We retain c if $-\log_{10}(p_c) > \tau_{\text{feat}}$. Let \mathcal{R}^{sel} denote the retained set; all subsequent steps use \mathbf{x} restricted to \mathcal{R}^{sel} .

Scoring. For samples i and j , define a dissimilarity d_{ij} using one of: (i) $\|\mathbf{x} - \mathbf{y}\|_2$, (ii) $1 - \rho_{\text{Pearson}}(\mathbf{x}, \mathbf{y})$, (iii) $1 - \rho_{\text{Spearman}}(\mathbf{x}, \mathbf{y})$. For each dissimilarity, we fit separate univariate Gaussian mixture models to the empirical distance distributions of positive (same-identity) and negative (different-identity) pairs, obtaining densities $p_{\text{pos}}(d)$ and $p_{\text{neg}}(d)$. The match score is the likelihood ratio

$$\Lambda(d_{ij}) = \frac{p_{\text{pos}}(t)}{p_{\text{neg}}(t)} \Big|_{t=d_{ij}},$$

and we declare a match if $\Lambda(d_{ij}) > \tau_{\Lambda}$.

3.3.2. Hand-engineered Descriptors

We form a single-channel image by summing all m/z intensities at each spatial location. From this image, we extract three classic grayscale descriptors: (i) local binary patterns (LBP) [18], (ii) histogram of oriented gradients (HOG) [7], and (iii) scale-invariant feature transform (SIFT) [17]. Each descriptor yields a feature vector, which is ℓ_2 -normalized and used for verification.

3.3.3. Convolutional Neural Networks

We implement the embedding function f using six CNNs – five pretrained and one custom-designed network.

Off-the-shelf CNN Backbones for Feature Extraction.

We use five pretrained backbones: *ResNet-18* [10], *ResNet-101* [10], *VGG-16* [32], *Inception-v3* [34], and *GoogLeNet* [34], followed by a lightweight projection layer forming embedding function f .

Custom-designed CNN. We design a custom convolutional network architecture composed from three modules: (i) a convolutional backbone, (ii) a downsampling stack, and (iii) a residual multi-layer perceptron head.

We choose an anisotropic 2×8 kernel in (i), reflecting the structure of the input data (497×2000) and also the fact that

peaks in GC×GC–ToF-MS typically appear as elongated blobs stretched along the second (retention-time) dimension (see Fig. 1), so a wider kernel in this direction provides a better inductive bias for capturing peak shapes.

A detailed description is in the Supplementary Sec. S1.5.

3.4. Identity Verification

Let $f : \mathbb{R}^{C \times H \times W} \rightarrow \mathbb{R}^D$ be an embedding function, and let $d : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}_{\geq 0}$ be a dissimilarity (e.g., Euclidean or $d_{\text{cos}}(\mathbf{x}, \mathbf{y}) := 1 - \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\|_2 \|\mathbf{y}\|_2}$). For a pair $(\mathbf{C}_i, \mathbf{C}_j)$, define

$$d_{ij} := d(f(\mathbf{C}_i), f(\mathbf{C}_j)).$$

At test time, we declare “same identity” if $d_{ij} < \tau$.

Metric Learning with Triplet Loss. We adopt the metric-learning setting with the triplet loss [30] for training f . We assume an anchor sample a with a matched positive p from the same identity, and a negative n from a different identity.

We train the feature-embedding function f implemented as neural network to minimize the margin-based triplet loss

$$\mathcal{L}(\mathbf{C}_i, \mathbf{C}_j; f, d, \gamma) = \frac{1}{|\mathcal{T}|} \sum_{(a,p,n) \in \mathcal{T}} \max\{d_{ap} - d_{an} + \gamma, 0\} \quad (4)$$

where $\gamma > 0$ is a predefined margin, $\gamma = 1$ in our experiments. The hinge term vanishes when the margin constraint $d_{ap} + \gamma \leq d_{an}$ is already satisfied, encouraging tight same-identity clusters and well-separated different-identity clusters in the embedding space.

Implementation Details. We implemented all methods in PyTorch [19] and trained neural networks using the AdamW [16] optimizer with fixed learning rate $3 \cdot 10^{-5}$. Standard techniques were used to prevent overfitting (see Supplementary Sec. S1.9 for a detailed description). The training was performed on a cluster of NVIDIA A100 GPUs with 40 GB of RAM and on a cluster of AMD MI250x GPUs with 64 GB of RAM.

4. Human Scent Dataset

The Human Scent Dataset (HSD) is a curated dataset consisting of 504 samples from 40 identities from a previously published 4D-C dataset [11], and additional newly collected 2,024 samples from 212 identities. In total, HSD contains 129 male identities contributing 1,178 samples and 123 female identities contributing 1350 samples. Data were acquired using the Pegasus® 4D-C and a higher resolution Pegasus® BT 4D, which provides enhanced mass-to-charge ratio precision. Sensor instrumentation is in the Supplementary Sec. S1.1, with dataset statistics and per-identity counts (by sex) in the Supplementary Table S1 and Fig. S1.

Data Acquisition. Chromatographic measurements require a significant time investment. Collecting each human scent sample – including all preparation steps – takes ≈ 20 minutes. Sensitivity and specificity enhancement takes ≈ 4 hours. The measurement and cleaning phase lasts about 2.5 hours, with data export requiring up to 4 hours depending on computational resources. GC \times GC–ToF–MS data files average ≈ 3 GB each. In total, our HSD dataset holds a modest number of samples compared to contemporary datasets but requires ≈ 7.5 TB of storage.

Although the current protocol is not yet suitable for large-scale or real-time use, ongoing work is already reducing the measurement time. Early tests of non-contact sampling reduce preparation effort and contamination risk. Advances in automation, calibration, and streamlined data processing further promise faster measurement and export.

All participants provided their informed consent.

5. Experiments

In all our experiments, we use identity- and gender-aware 10-fold cross-validation on the HSD dataset with fixed splits. First, we introduce evaluation metrics, then experiments on channel representations and identity verification.

5.1. Evaluation Metrics

5.1.1. Channel Representation Learning

Model. For a pixel \mathbf{u} with spectrum $\{(m_{n,\mathbf{u}}, I_{n,\mathbf{u}})\}_{n=1}^{N(\mathbf{u})}$, we approximate intensities with kernel basis κ and $\mathbf{C}_{k,\mathbf{u}}$

$$\hat{I}_{n,\mathbf{u}} = \sum_{k=1}^C \mathbf{C}_{k,\mathbf{u}} \kappa(m_{n,\mathbf{u}}; \theta_k).$$

Objective. Parameters $\{\alpha_k, \theta_k\}_{k=1}^C$ are learned by minimizing the weighted MSE over training pixels Ω :

$$E = \sum_{\mathbf{u} \in \Omega} \sum_{n=1}^{N(\mathbf{u})} \frac{1}{w_{n,\mathbf{u}}} (I_{n,\mathbf{u}} - \hat{I}_{n,\mathbf{u}})^2.$$

We use either uniform $w_{n,\mathbf{u}} = 1$ or inverse-intensity

$$w_{n,\mathbf{u}} = \max(I_{\Omega}/100, I_{n,\mathbf{u}}), \quad I_{\Omega} = \max_{\mathbf{u} \in \Omega, n} I_{n,\mathbf{u}}.$$

After training, $\{\alpha_k, \theta_k\}$ are frozen and reused for all pixels.

5.1.2. Metric Learning

We report two metrics, the area under the ROC curve (AUROC) and TPR@C-FPR. Both use the set of dissimilarities $\{d_{ij}\}$ defined in Sec. 3.4 and ground-truth binary labels $\{y_{ij}\}$. If two samples \mathcal{S}_i and \mathcal{S}_j belong to the same identity $y_{ij} = 1$, $y_{ij} = 0$ otherwise. We compute d_{ij} and declare the same identity ($\hat{y}_{ij} = 1$) iff $d_{ij} < \tau$; otherwise different ($\hat{y}_{ij} = 0$).

Definitions. Given the set of positive pairs \mathcal{P} and the set of negative pairs \mathcal{N} (with cardinalities $|\mathcal{P}|$ and $|\mathcal{N}|$), let

$$\text{TP}(\tau) = |\{(\mathcal{S}_i, \mathcal{S}_j) \in \mathcal{P} : d_{ij} < \tau\}|, \text{ and} \quad (5)$$

$$\text{FP}(\tau) = |\{(\mathcal{S}_i, \mathcal{S}_j) \in \mathcal{N} : d_{ij} < \tau\}|. \quad (6)$$

The true positive rate (TPR) and the false positive rate (FPR) at threshold τ are:

$$\text{TPR}(\tau) = \frac{\text{TP}(\tau)}{|\mathcal{P}|}, \quad \text{FPR}(\tau) = \frac{\text{FP}(\tau)}{|\mathcal{N}|}.$$

Area under ROC curve (AUROC). The receiver operating characteristic (ROC) curve is a plot of $\text{TPR}(\tau)$ versus $\text{FPR}(\tau)$ as τ varies over the range of observed dissimilarities. The AUROC summarizes this curve into a single scalar by integrating TPR with respect to FPR:

$$\text{AUROC} = \int_0^1 \text{TPR}(\text{FPR}^{-1}(u)) du.$$

An AUROC of 1.0 indicates perfect separability, while 0.5 corresponds to random guessing.

TPR@C-FPR. We measure the TPR at the threshold τ_c that ensures $\text{FPR}(\tau_c) \leq c_{\text{FPR}}$. Formally,

$$\tau_c = \min\{\tau \mid \text{FPR}(\tau) \leq c_{\text{FPR}}\},$$

then

$$\text{TPR@C-FPR} = \text{TPR}(\tau_c).$$

This metric captures how well the model correctly identifies positive pairs under a strict upper bound on false positives.

5.2. Channel Representation Learning

We apply a 10-fold cross-validation on the dataset of 2,528 samples by 497×2000 retention times by “number of spectral samples”. We train the Gaussian kernel (Sec. 3.2) by minimizing the weighted reconstruction loss E (Eq. 5.1.1) with stochastic gradient descent. Kernel centers $\{\mu_k\}_{k=1}^C$ are fixed by a simple, m/z -aware schedule: place $(C - 2)$ centers uniformly at interval midpoints over $[0, 300]$, and place the remaining two at the midpoints of an equal two-way partition of $[300, 800]$. This layout concentrates capacity where mass spectra carry most signal (low m/z) while preserving coverage of higher m/z values. Kernel widths σ_k and amplitudes α_k are learned during training.

We evaluate two weighting schemes in the reconstruction loss (Eq. 5.1.1): uniform and inverse-intensity (per- m/z weight $\propto 1/\text{ion count}$). Both kernels and schemes achieve $\text{MSE} < \varepsilon$ for all C , with $\varepsilon = 3 \cdot 10^{-4}$, indicating high-fidelity reconstructions. Per-setting results are provided for completeness in the Supplementary Table S3 (omitted from the main paper).

5.3. Feature Embedding Experiments

A detailed evaluation of four approaches to embedding of GC×GC–ToF-MS channel representations is presented in the following sections.

5.3.1. Baseline Method Developed by Chemists

We implemented the method proposed in [6] (see Sec. 3.3.1) and evaluated the method on the 4D-C dataset [11] only, since the method of [6] requires annotated compound quantities not available in the HSD dataset. We benchmark [6] against our custom-designed CNN, which operates on a single-channel input – the sum of the m/z intensities at each spatial location. Two compound selection methods were used to build the vector of compound quantities: (i) selection of all compounds, and (ii) application of the Wilcoxon statistical test [35] with selection thresholds $\tau_{\text{feat}} \in \{0.1, 0.5\}$, as recommended in [6].

Discussion. The method yields competitive 10-fold cross-validation results of AUROC = 79.3% and TPR@5%-FPR = 42.0% (see Table 1), surpassing our neural-network method on the 4D-C dataset. We attribute this outcome to the limited dataset size – there were only four identities in the validation portion of each fold. Note that in contrast to our custom-designed CNN that operates on single-channel inputs, the analytical-chemists baseline method uses all available mass spectra to detect chemical compounds and their quantities in the sample, later used for verification. For detailed results, see Suppl. Sec. S1.6.

5.3.2. Hand-engineered Descriptors

We benchmark three hand-engineered descriptors: (i) LBP [18], (ii) HOG [7], and (iii) SIFT [17]. We convert each GC×GC–ToF-MS measurement to its single-channel representation by summing the intensity values of all m/z ratios in each pixel. The similarity of produced descriptors is measured by cosine similarity.

We comprehensively evaluated the three descriptors across a broad range of hyperparameters (see Appendix Sec S1.7 for detailed description of the tested hyperparameters). For complete results, see Supplementary Sec. S1.7

Discussion. The best-performing configuration used the uniform-mapping LBP with radius $r = 1$, yielding an average cross-validation AUROC = 78.1% (see Table 1). The highest TPR@5%FPR was achieved by HOG with a 16×16 -pixel cell, six orientation bins, and a 3×3 cell block, reaching TPR@5%FPR = 40.5. Our custom-designed CNN using single-channel input outperformed the descriptors by more than 10% in both AUROC and TPR@5%FPR.

5.3.3. Convolutional Neural Networks

In CNN experiments, we evaluated our custom CNN and five pretrained backbones: *ResNet-18* [10], *ResNet-*

101 [10], *VGG-16* [32], *Inception-v3* [34], and *GoogLeNet* [34], followed by a lightweight projection layer forming the embedding function. The weights of the pretrained backbones were fixed except for the first convolution and the appended projection layer. We minimized the triplet loss (3). The dissimilarity $d_{i,j}$ was d_{\cos} . We used the channel representations with Gaussian kernels κ_G (2) and channel number $C = 3$ for the pretrained backbones. For our custom CNN, we evaluate channel counts $C \in \{1, 3, 5, 10, 20\}$. The choices $C \in \{1, 3, 5\}$ follow analytical practice: known compounds are often detected by monitoring very large ion counts at ≤ 5 selected m/z values (sometimes a single m/z). We include $C \in \{10, 20\}$ to quantify the effect of additional channels under a constrained computational budget.

The dimension of the vector embedding $D = 4096$. During training, we used the hard-negative mining strategy – in each epoch and for each identity, we uniformly selected one anchor sample, and we retained only the hardest 20% of the positives and the same number of the hardest negatives.

Discussion. In Table 1, we report the best-performing CNNs along with the best results of hand-engineered descriptors and the analytical-chemistry baseline. Although our custom-designed CNN achieves the best AUROC, the performance of CNNs is overall very similar, with AUROC near 90% and TPR@5%FPR around 50%. In contrast to a standard approach of chemical analysis methods that requires aligning the GC×GC–ToF-MS measurements into a common reference frame, our experiments show that on our HSD and for modern CV embeddings, canonical registration did not improve verification. Robustness to stronger t_1/t_2 drift and chemically interpretable tasks remains to be tested.

We argue that learning on unaligned GC×GC–ToF-MS improves identity verification because: (i) it marginalizes over run-specific drifts instead of relying on noisy alignment, (ii) alignment is a lossy bottleneck that discards discriminative chromatographic/spectral and retention time-mass spectral coupling cues, and (iii) estimated warps amplify noise and inject artifacts.

Surprisingly, when only a single channel representation kernel is used, our custom-designed CNN achieves an impressive AUROC of 85.8%. We believe that the high performance of our CNN is caused by the rectangular shape of convolutions – the second dimension of a GC×GC–ToF-MS measurement is constructed by modulating the analysis time, which makes the responses of compounds significantly prolonged in the y-axis, and rectangular convolutional kernels capture those responses better.

The complete results with varying channel representation kernels and representation learning weightings are given in the Supplementary Sec. S1.10.

C	no registration		$c_{\cos}(\bar{\mathbf{u}}_c, \bar{\mathbf{s}}_c)$				FCN(\mathbf{u}, \mathbf{s}) $^+_{\cos}$				
			$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$		$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$		
	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	
baseline	–	79.3±6.2	42.0±12.6	–	–	–	–	–	–	–	
custom CNN	1	75.5±7.2	29.2±14.0	75.2±7.6	29.7±16.2	75.5±7.6	26.6±13.1	76.8±7.7	33.2±15.4	76.4±7.9	30.8±14.2
HOG	1	73.0±6.6	39.1± 7.8	74.3±7.2	39.1± 7.8	74.4±7.3	39.3± 7.7	76.0±4.9	33.4± 7.2	76.7±5.0	33.8± 7.1
LBP	1	78.1±3.2	15.8± 2.2	77.1±2.8	15.8± 2.2	77.1±2.8	15.5± 2.1	76.8±4.0	22.4± 4.3	78.1±4.3	22.9± 4.9
SIFT	1	66.6±6.0	18.6± 5.3	63.8±4.1	15.5± 4.0	63.4±3.9	15.1± 3.8	59.0±6.7	13.4± 4.0	63.8±6.5	17.4± 4.9
custom CNN	1	85.8±5.1	39.0±11.8	88.0±3.1	48.3± 9.2	88.0±3.6	50.5±10.5	87.9±3.5	49.1± 8.9	87.4±2.9	50.2± 9.1
VGG16	3	77.0±4.6	40.5± 6.6	79.5±9.6	42.9±10.4	79.8±9.4	42.2± 9.4	74.6±7.2	26.2± 8.3	74.5±6.9	26.6± 6.6
ResNet-18	3	87.5±3.9	44.9± 8.8	87.3±3.5	42.9± 5.2	86.9±3.0	41.1± 5.9	85.2±3.8	36.1± 7.1	85.5±4.3	36.2± 8.2
ResNet-101	3	86.3±4.1	41.5±11.4	86.0±4.2	38.5± 8.7	86.5±4.1	40.6± 7.7	84.5±4.7	35.3± 8.7	84.9±4.4	35.2± 9.1
GoogLeNet	3	87.9±3.5	46.8± 7.7	86.8±3.9	44.3± 8.4	86.7±3.4	44.7± 7.5	85.0±3.8	35.2± 8.9	84.9±4.1	36.1±10.3
Inception-v3	3	86.0±3.5	40.2± 5.6	85.5±3.3	39.2± 5.0	85.5±3.5	40.2± 5.1	83.5±3.9	32.8± 6.8	84.0±4.0	34.3± 9.1
custom CNN	3	88.3±3.8	52.4± 9.3	88.2±2.9	50.8± 5.1	87.3±3.5	48.6± 7.1	84.6±4.5	42.9± 6.8	83.8±5.1	40.0± 7.8
custom CNN	5	89.5±3.4	53.3± 8.3	88.3±3.4	51.3± 6.5	88.4±3.4	50.7± 6.7	85.8±3.4	46.5± 5.5	85.4±3.8	43.9± 6.2
custom CNN	10	88.7±3.9	53.7± 9.7	87.0±4.1	43.2± 8.4	86.3±3.9	50.0± 6.8	83.0±4.7	36.8± 9.8	83.3±4.9	39.5± 8.3
custom CNN	20	89.0±3.7	53.1± 7.1	87.5±3.9	49.0± 6.6	87.9±3.0	50.3± 6.6	84.9±3.9	43.6± 5.6	84.2±4.8	43.4± 7.1

Table 1. Best cross-validation Area Under ROC Curve and TPR@5%FPR for five spatial-alignment strategies on channel representation (Sec. 3.2) of 2D Gas Chromatography Mass Spectrometry measurements. Top: 4D-C dataset [11] with manual annotations of chemical compounds, bottom: Human Scent Dataset (Sec. 4). Results reported as $\mu \pm \sigma$ across 10 folds. Identity verification performed with: (i) analyst chemists *baseline* [6], (ii) deterministic descriptors LBP [18], HOG [7], and SIFT [17], (iii) custom-designed Convolutional Neural Network (CNN), and (iv) five pretrained neural network backbones *ResNet-18* [10], *ResNet-101* [10], *VGG-16* [32], *Inception-v3* [34], and *GoogLeNet* [34] each with one appended and trained projection layer. Number of channel representation kernels $C \in \{1, 3, 5, 10, 20\}$ (not applicable to baseline chemistry pipeline). Analyst chemistry baseline uses all available mass spectra \mathcal{S} instead of channel representation. We perform registration using the weighted-cosine $c_{\cos}(\bar{\mathbf{u}}_c, \bar{\mathbf{s}}_c)$ and fully convolutional network-based FCN(\mathbf{u}, \mathbf{s}) $^+_{\cos}$ compound-peak detectors introduced in [11], evaluated on the $\mathcal{R}^{(22)}$ and $\mathcal{R}^{(24)}$ sets (22 and 24 compounds). For detailed results, refer to Supplementary Tables S4, S5, and S8.

In Supplementary Sec. S1.8, we compare our custom CNN on high- and low-resolution subsets of HSD.

6. Limitations

Although our results establish feasibility, the following limitations apply:

- (i) accuracy ($\approx 53\%$ TPR5% FPR; AUROC $\approx 89.5\%$) well below that of established modalities;
- (ii) external validity across labs, climates, diets, hygiene, medications, and environments, remains to be investigated;
- (iii) long-term permanence and drift are unquantified;
- (iv) the observation that explicit canonical registration is unnecessary is currently demonstrated only on Human Scent Dataset, with robustness under stronger retention-time drift and in settings requiring chemical interpretability still to be evaluated.

7. Conclusion

In this work, we systematically evaluated ten approaches for odor-based identity verification using raw GC×GC–ToF-MS data, represented as multi-channel im-

ages. In contrast to common analytical pipelines, the best results on our dataset are achieved without spatial alignment, indicating that modern learned embeddings can tolerate the observed retention-time drift.

Our experiments on a newly proposed dataset of 252 individuals demonstrate the feasibility of scent-based recognition: despite modest initial performance compared to established biometrics, the best performing approach achieves an $\approx 53\%$ true positive rate at a 5% false positive rate with an 89.5% AUROC.

These findings confirm that odor signatures can serve as a biometric trait when processed with modern feature embedding techniques. At the same time, the results underscore clear opportunities for improvement, including more expressive representations and more efficient measurement protocols.

Acknowledgements. The research reported in this paper has been partly funded by BMK, BMAW, and the State of Upper Austria in the frame of the SCCH competence center INTEGRATE [Project 1.6 TFI (Transferable Intelligence)] part of the FFG COMET Competence Centers for Excellent Technologies Programme, and by the Czech Technical University in Prague grant No. SGS23/173/OHK3/3T/13, and by the Ministry of the Interior of the Czech Republic grant No. VJ01010123, and by the Ministry of Education, Youth and Sports of the Czech Republic through the e-INFRA CZ (ID:90254).

References

- [1] Julián Eduardo Ballén Castiblanco, Victor Hugo Calvacanti Ferreira, Carlos Alberto Teixeira, and Leandro Wang Hantao. Classification of produced water samples using class-oriented chemometrics and comprehensive two-dimensional gas chromatography coupled to mass spectrometry. *Talanta*, 268:125343, 2024. [3](#)
- [2] Stephanie Brener, Kobi Snitz, and Noam Sobel. An electronic nose can identify humans by the smell of their ear. *Chemical Senses*, 49, 2024. [2](#)
- [3] Bogusław Buszewski, Damian Grzywinski, Tomasz Ligor, Tadeusz Stacewicz, Zygmunt Bielecki, and Jacek Wojtas. Detection of Volatile Organic Compounds as Biomarkers in Breath Analysis by Different Analytical Techniques. *Bioanalysis*, 5(18):2287–2306, 2013. [3](#)
- [4] Andrea Caratti, Simone Squara, Carlo Bicchi, Erica Liberto, Marco Vincenti, Stephen E. Reichenbach, Qingping Tao, Daniel Geschwender, Eugenio Alladio, and Chiara Cordero. Boosting comprehensive two-dimensional chromatography with artificial intelligence: Application to food-omics. *TrAC Trends in Analytical Chemistry*, 174:117669, 2024. [3](#)
- [5] Lauren J. Colón-Crespo, Danay Herrera-Hernández, Howard Holness, and Kenneth G. Furton. Determination of VOC marker combinations for the classification of individuals by gender and race/ethnicity. *Forensic Science International*, 270:193–199, 2017. [3](#)
- [6] Vincent Cuzuel, Roman Leconte, Guillaume Cognon, Didier Thiébaud, Jérôme Vial, Charles Sauleau, and Isabelle Rivals. Human odor and forensics: Towards Bayesian suspect identification using GC×GC–MS characterization of hand odor. *Journal of Chromatography B*, 1092:379–385, 2018. [3](#), [5](#), [7](#), [8](#), [11](#), [13](#)
- [7] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, pages 886–893. Ieee, 2005. [4](#), [5](#), [7](#), [8](#)
- [8] Victor H. C. Ferreira, Leandro W. Hantao, and Ronei J. Poppi. Use of color based chromatographic images obtained from comprehensive two-dimensional gas chromatography in authentication analyses. *Talanta*, 234:122616, 2021. [3](#)
- [9] Chantrell J. G. Frazier, Vidia A. Gokool, Howard K. Holness, DeEtta K. Mills, and Kenneth G. Furton. Multivariate regression modelling for gender prediction using volatile organic compounds from hand odor profiles via HS-SPME-GC-MS. *PLOS ONE*, 18(7), 2023. [3](#)
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016. [5](#), [7](#), [8](#), [17](#)
- [11] Jan Hlavsa, Radim Spetlík, Jana Čechová, Petra Pojmanová, Jiří Matas, and Štěpán Urban. Sex Classification from Human Scent Using Image Interpretation of 2D Gas Chromatography-Mass Spectrometry Data. In *Image Analysis*, pages 457–470. Springer Nature Switzerland, 2025. [3](#), [4](#), [5](#), [7](#), [8](#), [11](#), [12](#), [13](#), [15](#), [17](#)
- [12] Chaiyanut Jirayupat, Kazuki Nagashima, Takuro Hosomi, Tsunaki Takahashi, Benjarong Samransuksamer, Yosuke Hanai, Atsuo Nakao, Masaya Nakatani, Jianguang Liu, Guozhu Zhang, Wataru Tanaka, Masaki Kanai, Takao Yasui, Yoshinobu Baba, and Takeshi Yanagida. Breath odor-based individual authentication by an artificial olfactory sensor system and machine learning. *Chemical Communications*, 58(44):6377–6380, 2022. [1](#), [3](#)
- [13] Oleksii Kaminskyi, Jana Čechová, Ulrika Malá, Patrik Kania, Nikola Ladislavová, Petra Pojmanová, and Štěpán Urban. Device for Non-contact Odor Sampling. *Forenzní vědy, právo, kriminalistika*, 8(1):33–41, 2023. [2](#)
- [14] Hiroyuki Kataoka, Keita Saito, Hisato Kato, and Kazufumi Masuda. Noninvasive analysis of volatile biomarkers in human emanations for health and early disease diagnosis. *Bioanalysis*, 5(11):1443–1459, 2013. [3](#)
- [15] Nikola Ladislavová, Petra Pojmanová, Pavel Vrbka, Jana Šnupárková, and Štěpán Urban. Human scent signature on cartridge case survives gun being fired: A preliminary study on a potential of scent residues as an identification tool. *PLOS ONE*, 18(3):e0283259, 2023. [1](#), [3](#)
- [16] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *7th International Conference on Learning Representations*, 2019. [5](#)
- [17] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004. [4](#), [5](#), [7](#), [8](#)
- [18] Timo Ojala, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7):971–987, 2002. [4](#), [5](#), [7](#), [8](#), [14](#)
- [19] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, 2019. [5](#)
- [20] Dustin J Penn, Elisabeth Oberzaucher, Karl Grammer, Gottfried Fischer, Helena A Soini, Donald Wiesler, Milos V Novotny, Sarah J Dixon, Yun Xu, and Richard G Breerton. Individual and gender fingerprints in human body odour. *Journal of The Royal Society Interface*, 4(13):331–340, 2006. [1](#), [2](#), [3](#)
- [21] Ruud Peters, Rick Veenstra, Karin Heutinck, Albert Baas, Sandra Munniks, and Jaap Knotter. Human scent characterization: A review. *Forensic Science International*, 349:111743, 2023. [1](#)
- [22] Ludvík Pinc, Luděk Bartoš, Alice Reslová, and Radim Kotrba. Dogs Discriminate Identical Twins. *PLoS ONE*, 6(6):e20704, 2011. [1](#), [2](#)
- [23] Ludvik Pinc, Ludek Bartos, and Petra Vypelova. Resistance of human odours to extremely high temperature as revealed by trained dogs. *Czech Journal of Animal Science*, 61(4):172–176, 2016. [1](#)

- [24] Petra Pojmanová. *Molekulová skladba pachové signatury*. PhD thesis, VŠCHT, 2020. [1](#), [3](#)
- [25] Petra Pojmanová, Nikola Ladislavová, Veronika Škeříková, Patrik Kania, and Štěpán Urban. Human scent samples for chemical analysis. *Chemical Papers*, 74(5):1383–1393, 2020. [11](#)
- [26] Petra Pojmanová, Nikola Ladislavová, and Štěpán Urban. Development of a Method for the Measurement of Human Scent Samples Using Comprehensive Two-Dimensional Gas Chromatography with Mass Detection. *Separations*, 8(12):232, 2021. [11](#)
- [27] Petra Pojmanová, Nikola Ladislavová, and Štěpán Urban. Human Skin Scent Class and Individual Identification, 2022. [1](#), [3](#)
- [28] Katharine A. Prokop-Prigge, Kathryn Greene, Lauren Varallo, Charles J. Wysocki, and George Preti. The effect of ethnicity on human axillary odorant production. *Journal of Chemical Ecology*, 42(1):33–39, 2016. [3](#)
- [29] Irene Rodriguez-Lujan, Gonzalo Bailador, Carmen Sanchez-Avila, Ana Herrero, and Guillermo Vidal-de Miguel. Analysis of pattern recognition and dimensionality reduction techniques for odor biometrics. *Knowledge-Based Systems*, 52:279–289, 2013. [1](#), [2](#), [3](#)
- [30] Florian Schroff, Dmitry Kalenichenko, and James Philbin. FaceNet: A unified embedding for face recognition and clustering. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 815–823, 2015. [5](#)
- [31] Mika Shirasu and Kazushige Touhara. The scent of disease: volatile organic compounds of the human body related to disease and disorder. *The Journal of Biochemistry*, 150(3):257–266, 2011. [3](#)
- [32] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. In *Proceedings of 3rd International Conference on Learning Representations*, 2015. [5](#), [7](#), [8](#), [17](#)
- [33] Rex A Stockham, Dennis L Slavin, and William Kift. Survivability of human scent. *Forensic Science Communications*, 6(4), 2004. Publisher: Federal Bureau of Investigation. [1](#)
- [34] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015. [5](#), [7](#), [8](#), [17](#)
- [35] Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945. [7](#)

S1. Supplementary

In the following sections, we describe the instrumentation of GC×GC–ToF–MS sensor, the compound peak detectors introduced in [11], including a detailed overview of our custom CNN architecture. We also present comprehensive results for the baseline method developed by analytical chemists [6], and for the deterministic descriptors.

S1.1. Sensor Instrumentation

Analyses were conducted on an Agilent 7890B gas chromatograph (Agilent, Santa Clara, CA, USA) coupled to time-of-flight mass spectrometers Pegasus[®] 4D-C and Pegasus[®] BT-4D (LECO Corp., St. Joseph, MI, USA). Automated sample introduction was performed with a MultiPurpose Sampler MPS Robotic (Gerstel, Mülheim, Germany). Raw data were exported from the ChromaTOF[®] software (versions 4.72.0.0, 5.51.06.0, and 5.55.41_BT; LECO Corp., St. Joseph, MI, USA).

The GC×GC–ToF–MS configuration followed a previously reported layout [26] employing a reverse column arrangement: a medium-polar primary column Rtx-200MS (Restek, Bellefonte, PA, USA, 30 m × 0.25 mm × 0.25 μm) and a non-polar secondary column TG-5HT (Thermo Fisher Scientific, Waltham, MA, USA, 1 m × 0.25 mm × 0.25 μm). Of the total secondary-column length, 10 cm were located in the modulator, 69 cm in the secondary oven, and 21 cm in the GC-MS transfer line. A 2 m pre-column of the same stationary phase as the primary column was installed. When using the Pegasus BT-4D detector, the secondary column totaled 1.1 m to accommodate 31 cm within the GC-MS transfer line.

Prior to injection, samples were homogenized by shaking for 10 min followed by 10 min in an ultrasonic bath. Injections of 1 μL were made with a 1 μL air plug in the syringe, at 280°C in splitless mode (2 min splitless time). Carrier gas was operated at a constant-flow rate of 1.5 mL/min. The oven program was: 40°C (2 min), ramp at 5°C/min to 320°C (10 min), for a total runtime of 68 min (4080 s). The secondary oven was maintained at +5°C relative to the primary oven and the modulator at +20°C relative to the primary. Cryogenic modulation of the secondary column used dried air at -80°C. The run comprised three modulation segments: 500-1704 s (period 6 s; hot pulse 1.8 s; cold pulse 1.2 s), 1704-2592 s (period 8 s; hot 2.4 s; cold 1.6 s), and 2592-4080 s (period 10 s; hot 3.3 s; cold 1.7 s).

The transfer line was held at 280°C and the ion source at 250°C. Electron ionization (EI) at 70 eV was used. Total ion current (TIC) data were collected over m/z 29–800 at 200 spectra s^{-1} . The acquisition delay was 500 s, and the detector voltage was set 200 V above the tuned value to enhance sensitivity.

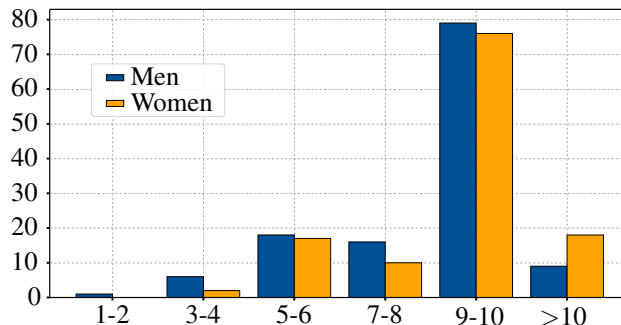


Figure S1. Histogram of measurement counts per identity in the HSD dataset, grouped by biological sex (see Sec. 4). X-axis: measurements per identity, y-axis: number of subjects.

Dataset	Total num.	n	Number of samples per identity			
			min	max	avg	median
4D-C	504	40	10	37	12.6	10
HSD	2528	252	1	76	10.0	10

Table S1. Dataset statistics. From left to right: dataset name, total number of samples, number of identities n in the dataset, and: (i) minimum, (ii) maximum, (iii) average, and (iv) median number of measurements per identity.

S1.2. Compound Peak Detectors

We locate the characteristic peak of each compound using an annotated 4D-C dataset of 40 individuals (504 total samples) from [11, 25]. Although we train our compound-peak detectors on a small annotated set, we subsequently rectify the entire Human Scent Dataset (Sec. 4) using these learned detectors, eliminating the need for extensive annotation by chemical experts. To align the spectral resolutions, we quantized the high-spectral-resolution measurements to match the dimensions of the low-spectral-resolution data. We then stored the detected peak positions and applied the registration procedure to each channel representation using these stored positions.

S1.2.1. Chemical Compounds Being Detected

In the registration step, we followed [11] exactly. We used the set of 22 chemical compounds from the paper and also experimented with the addition of two more compounds, Heptacosane and Nonacosane. See the list of compounds in Table S2. Note that a standardized mass spectrum of all listed compounds is available in public databases.

S1.2.2. Bounding Box Constraint

Many compounds share similar mass spectra (see Fig. 1), so *positional cues* are crucial as discussed in [11]. We restrict the search for compound c to a bounding box around its

#	Compound
1	Tetradecane, 1-chloro-
2	2,5-pyrrolidinedione, 3-decyl-
3	2,5-pyrrolidinedione, 3-dodecyl-
4	2,5-pyrrolidinedione, 3-tetradecyl-
5	2,5-pyrrolidinedione, 3-hexadecyl-
6	Squalene
7	Cholesterol
8	Cholesta-3,5-dien-7-one
9	Cholest-4-en-3-one
10	Cholesta-4,6-dien-3-one
11	Hexadecanoic acid, dodecyl ester
12	9-Hexadecenoic acid, tetradecyl ester
13	Hexadecanoic acid, tetradecyl ester
14	9-Hexadecenoic acid, hexadecyl ester
15	Hexadecanoic acid, hexadecyl ester
16	9-Hexadecenoic acid, octadecyl ester
17	Hexadecanoic acid, octadecyl ester
18	1-Dodecanol
19	Dodecane, 1-chloro-
20	Dodecanamide, N-allyl-
21	Fumaric acid, 2-dimethylaminoethyl heptyl ester
22	Cholesta-3,5-diene (isomer 2)
23	Heptacosane
24	Nonacosane

Table S2. List of compounds used for registration in our experiments, following [11]. The first 22 compounds correspond to the set actually used for registration in [11]. Heptacosane and Nonacosane were only mentioned in the original paper but not used there; in our experiments, we evaluated both the 22-compound and 24-compound variants.

average position $\bar{\mathbf{u}}_c$:

$$\text{BBox}(\bar{\mathbf{u}}_c, w, h) = \left\{ \mathbf{u} \in \Omega \mid |\bar{u}_{1,c} - u_1| < \frac{w}{2}, \quad |\bar{u}_{2,c} - u_2| < \frac{h}{2} \right\}, \quad (\text{S1})$$

$$(\text{S2})$$

where $\Omega := \{1, \dots, H\} \times \{1, \dots, W\}$ is the pixel grid, and w, h are the bounding-box dimensions.

S1.2.3. Weighted Cosine Peak Detector

Cosine similarity alone can yield spurious matches. To reduce those, we apply a spatial penalty that was presented in [11]:

$$\hat{\mathbf{u}}_c^{c_{\text{cos}}(\bar{\mathbf{u}}_c, \bar{s}_c)} = \arg \max_{\mathbf{u} \in \text{BBox}(\bar{\mathbf{u}}_c)} \left[\cos(\bar{s}_c, s_{\mathbf{u}}) - p(\bar{\mathbf{u}}_c, \mathbf{u}) \right],$$

where

$$p(\bar{\mathbf{u}}_c, \mathbf{u}) = (|\bar{u}_{1,c} - u_1| + 6|\bar{u}_{2,c} - u_2|)^4.$$

The exponent $(\cdot)^4$ strongly penalizes large deviations, and the factor of 6 accounts for typically wider shifts along the first chromatographic dimension. This detector thus prefers spectra near $\bar{\mathbf{u}}_c$ that match \bar{s}_c .

Cosine similarity measures the angular alignment between two spectra, making the approach invariant to absolute intensity scaling.

S1.2.4. Fully Convolutional Network (FCN) Peak Detector

A fully convolutional network (FCN) that was shown in [11] predicts *class likelihood* at each pixel. Let Ω_c be the set of pixels classified by the FCN as belonging to compound c . We combine the FCN prediction with cosine similarity:

$$\hat{\mathbf{u}}_c^{\text{FCN}(\mathbf{u}, \mathbf{s})_{\text{cos}}^+} = \arg \max_{\mathbf{u} \in \Omega_c} \left[\text{FCN}(\mathbf{u}, s) + \cos(\bar{s}_c, s_{\mathbf{u}}) \right].$$

In both detectors, the final peak location is the pixel \mathbf{u} that maximizes the respective score.

S1.3. Canonical Frame Registration

As noted in [11], “variations in instrument conditions (e.g., temperature changes, aging, injection timing) shift compound retention times and lead to misalignment of GC×GC–ToF–MS measurements.” Canonical-frame registration was therefore considered a key step in their processing pipeline. We reimplemented the procedure introduced in [11] and conducted experiments with and without registration to assess its impact in our setting.

Let $\mathcal{R}^{\text{ref}} = \{c_1, c_2, \dots\}$ denote the set of reference compounds. For each $c \in \mathcal{R}^{\text{ref}}$, let $\mathbf{u}_c \in \mathbb{R}^2$ be the reference location and s_c^{ref} the corresponding mass spectrum.

In each human scent sample (GC×GC–ToF–MS measurement) \mathcal{S} , we first detect the occurrences (see Sec. S1.2 for detectors description) of a fixed set of reference compounds (see Sec. S1.2.1 for the list). This set corresponds to the compounds that were reported as consistent across samples in [11]. For each compound $c \in \mathcal{R}^{\text{ref}}$ we obtain its detected position $\mathbf{u}_c^{(\mathcal{S})} \in \mathbb{R}^2$ and corresponding mass spectrum $s_c^{(\mathcal{S})}$.

Given the detected positions $\mathcal{U}^{(\mathcal{S})} = \{\mathbf{u}_c^{(\mathcal{S})} \mid c \in \mathcal{R}\}$ in sample \mathcal{S} , we construct a Delaunay triangulation between $\mathcal{U}^{(\mathcal{S})}$ and reference locations \mathcal{U} to establish correspondence for the registration step.

For each triangle in the Delaunay triangulation, we compute a local affine mapping that transforms the detected vertices $\mathbf{u}_{c_1}^{(\mathcal{S})}, \mathbf{u}_{c_2}^{(\mathcal{S})}, \mathbf{u}_{c_3}^{(\mathcal{S})}$ of sample \mathcal{S} to the corresponding reference vertices $\mathbf{u}_{c_1}, \mathbf{u}_{c_2}, \mathbf{u}_{c_3}$. A point $\mathbf{u}^{(\mathcal{S})}$ in the measured sample is first expressed in barycentric coordinates $(\lambda_1, \lambda_2, \lambda_3)$ with respect to $(\mathbf{u}_{c_1}^{(\mathcal{S})}, \mathbf{u}_{c_2}^{(\mathcal{S})}, \mathbf{u}_{c_3}^{(\mathcal{S})})$,

$$\mathbf{u}^{(\mathcal{S})} = \lambda_1 \mathbf{u}_{c_1}^{(\mathcal{S})} + \lambda_2 \mathbf{u}_{c_2}^{(\mathcal{S})} + \lambda_3 \mathbf{u}_{c_3}^{(\mathcal{S})}, \quad \lambda_1 + \lambda_2 + \lambda_3 = 1, \quad \lambda_i \geq 0.$$

The canonical position \mathbf{u} is then obtained by mapping the same barycentric coordinates to the reference triangle,

$$\mathbf{u} = \lambda_1 \mathbf{u}_{c_1} + \lambda_2 \mathbf{u}_{c_2} + \lambda_3 \mathbf{u}_{c_3}.$$

Finally, the entire human scent sample is interpolated on the transformed mesh using piecewise linear interpolation, yielding the canonical frame representation.

This process allows comparison of scent samples taken under different conditions as was introduced in [11].

S1.4. Triangular Kernel Function For Channel Representation

In addition to the Gaussian kernel function introduced in 3.2, we experimented with triangular kernel

$$\kappa_{\Delta}(m_n; \mu_k, w_k) = \max\left(1 - \frac{|m_n - \mu_k|}{w_k}, 0\right),$$

a piecewise linear function whose half-width $w_k > 0$ is learned. In this case, the kernel centers μ_k are set using the same heuristic as in Sec. 5.2.

We apply a 10-fold cross-validation on the dataset of 2528 samples by 497×2000 retention times by “number of spectral samples”. For the Triangular kernel, we again minimize the weighted reconstruction error E (5.1.1) using stochastic gradient descent and compare two weight schemes: uniform weighting and inversed-intensity weighting (5.1.1).

κ	w	$C = 3$	$C = 5$	$C = 10$	$C = 20$
G	U	$6.3 \cdot 10^{-8}$	$1.4 \cdot 10^{-7}$	$4.5 \cdot 10^{-7}$	$3.9 \cdot 10^{-8}$
G	I	$5.1 \cdot 10^{-6}$	$1.7 \cdot 10^{-5}$	$5.1 \cdot 10^{-7}$	$8.0 \cdot 10^{-7}$
Δ	U	$2.7 \cdot 10^{-5}$	$1.7 \cdot 10^{-4}$	$2.5 \cdot 10^{-5}$	$2.1 \cdot 10^{-5}$
Δ	I	$1.5 \cdot 10^{-4}$	$1.5 \cdot 10^{-4}$	$1.6 \cdot 10^{-4}$	$1.0 \cdot 10^{-6}$

Table S3. Mean squared reconstruction error of channel representation with C kernels over 10-fold cross-validation of Human Scent Dataset (Sec. 4) with two kernel types κ (Gaussian kernel G and Triangular kernel Δ) and two weight types (uniform U and inversed intensity I). Reporting only μ .

The results are presented in Tab. S3 The reconstruction errors of triangular kernels is \approx two orders of magnitude higher than the error of Gaussian kernels. We attribute this to the higher expressiveness and smoother approximation capability of the Gaussian kernel. Therefore, all results in the main paper are reported using Gaussian kernels.

S1.5. Architecture of Custom CNN

As a custom-designed feature extractor, we introduce a compact convolutional network, to realize the embedding function f . The model comprises three modules: (i) a convolutional backbone, (ii) a down-sampling stack, and (iii) a residual MLP head.

		Threshold τ_{feat}		
		not set	$\tau_{\text{feat}} = 0.1$	$\tau_{\text{feat}} = 0.5$
AUC	Euclidean	62.1 ± 6.5	49.6 ± 3.3	49.2 ± 2.6
	Pearson	57.7 ± 7.4	56.2 ± 6.7	55.2 ± 4.2
	Spearman	79.3 ± 6.2	56.1 ± 5.2	54.8 ± 2.9
TPR@5%FPR	Euclidean	13.7 ± 9.3	5.7 ± 2.0	3.7 ± 1.6
	Pearson	10.8 ± 9.7	9.8 ± 4.9	6.6 ± 2.1
	Spearman	42.0 ± 12.6	5.8 ± 3.5	1.4 ± 2.0

Table S4. Cross-validation Area Under ROC Curve and TPR@5%FPR of baseline method developed by analytical chemists [6] on 4D-C dataset [11].

Backbone consists of three stages, each featuring: (i) a 2D Convolution block with a kernel 2×8 , stride (1, 2), and padding (0, 1), (ii) a Residual 2D Convolution Block that enhances feature propagation, and (iii) dropout layers with probabilities scaled by the dropout strength (i.e., 0.1, 0.2 and 0.3).

Down-sampling Module applies four sequential 2D Convolution block layers using a 5×5 kernel, stride 2, and padding 1, each followed by dropout at $0.3 \times$ dropout strength to progressively reduce spatial dimensions.

Residual MLP Head maps the high-dimensional feature vector, the flattened down-sampled features, to the desired space representation, with an additional dropout of $0.5 \times$ dropout strength to mitigate overfitting.

(Residual) 2D Convolution Block sequentially applies a 2D convolution, 2D batch normalization, a ReLU activation, and dropout. In its residual variant, the block adds the input element-wise to its output. When the input and output channel dimensions differ, a depth-wise 2D convolution is optionally applied to the input to ensure proper dimensionality alignment.

S1.6. Evaluation of Baseline Method Developed by Analyst Chemists

The full results of the experiment from Sec. 5.3.1 are presented in Table S4.

S1.7. Evaluation of Deterministic Descriptors

In this section, we evaluate three deterministic feature descriptors, LBP, HOG, and SIFT, and systematically analyze how their hyperparameter configurations influence verification performance. Detailed results are reported in Table S5.

For LBP, we vary the neighbourhood radius, $r \in \{1, 2, 3\}$, and examine both the *basic* (default) and *uniform* mapping schemes. The resulting feature embedding vector is a normalized histogram over the LBP codes.

		parameters		no registration		$c_{\cos}(\bar{\mathbf{u}}_c, \bar{\mathbf{s}}_c)$				FCN(\mathbf{u}, \mathbf{s}) $^+_{\cos}$					
		θ_1	θ_2	θ_3	θ_4	$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$		$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$			
						AUC	TPR@5%	AUC	TPR@5%	AUC	TPR@5%	AUC	TPR@5%		
HOG		8x8	6	3x3		72.4 ± 7.3	39.6 ± 8.1	74.3 ± 7.2	39.1 ± 7.8	74.4 ± 7.3	39.3 ± 7.7	75.8 ± 4.3	33.0 ± 7.2	76.6 ± 4.5	33.1 ± 7.1
		8x8	6	2x2		73.0 ± 6.6	39.1 ± 7.8	73.4 ± 6.5	38.9 ± 7.6	73.5 ± 6.6	38.9 ± 7.7	75.3 ± 3.6	31.9 ± 6.8	76.2 ± 3.7	31.8 ± 7.1
		16x16	6	3x3		71.0 ± 8.1	40.5 ± 8.4	74.2 ± 7.8	39.8 ± 8.1	74.3 ± 7.9	39.9 ± 8.2	75.2 ± 4.5	33.0 ± 7.1	75.8 ± 4.8	32.8 ± 7.1
		8x8	9	3x3		72.2 ± 7.9	39.4 ± 8.0	73.9 ± 7.9	38.8 ± 8.0	74.0 ± 7.9	39.3 ± 8.1	76.0 ± 4.9	33.4 ± 7.2	76.7 ± 5.0	33.5 ± 7.1
		16x16	6	2x2		71.4 ± 7.6	39.1 ± 7.8	73.5 ± 7.5	39.1 ± 7.8	73.6 ± 7.5	39.2 ± 7.8	75.3 ± 4.3	32.0 ± 6.8	76.2 ± 4.5	31.7 ± 6.8
		8x8	9	2x2		72.9 ± 7.1	38.5 ± 7.5	73.3 ± 7.1	38.2 ± 7.5	73.3 ± 7.2	38.5 ± 7.8	75.7 ± 4.0	32.2 ± 6.8	76.5 ± 4.2	32.2 ± 6.9
		16x16	9	3x3		71.2 ± 8.5	40.1 ± 8.3	73.9 ± 8.2	39.5 ± 8.1	74.0 ± 8.3	39.6 ± 8.2	75.2 ± 4.9	33.1 ± 7.0	75.8 ± 5.2	33.2 ± 7.2
		8x8	12	2x2		72.6 ± 7.4	38.0 ± 7.5	73.1 ± 7.4	37.7 ± 7.7	73.2 ± 7.5	38.1 ± 7.8	75.7 ± 4.3	32.5 ± 6.8	76.6 ± 4.5	32.7 ± 6.8
		8x8	12	3x3		72.0 ± 8.2	39.1 ± 7.9	73.8 ± 8.1	38.6 ± 8.0	73.9 ± 8.2	39.0 ± 8.2	76.0 ± 5.1	33.8 ± 7.3	76.7 ± 5.3	33.8 ± 7.1
		16x16	9	2x2		71.8 ± 7.8	38.8 ± 7.7	73.5 ± 7.8	38.9 ± 7.9	73.5 ± 7.9	39.1 ± 8.0	75.6 ± 4.5	32.0 ± 6.8	76.4 ± 4.8	32.1 ± 6.7
		16x16	12	3x3		71.3 ± 8.6	39.9 ± 8.3	73.8 ± 8.3	39.3 ± 8.0	73.9 ± 8.4	39.5 ± 8.2	75.2 ± 5.1	33.5 ± 7.1	75.8 ± 5.3	33.5 ± 7.1
		16x16	12	2x2		71.8 ± 8.0	38.6 ± 7.7	73.5 ± 7.9	38.9 ± 7.9	73.5 ± 8.0	39.1 ± 8.0	75.6 ± 4.7	32.5 ± 7.0	76.4 ± 4.9	32.3 ± 6.7
LBP		1	d			78.1 ± 3.2	15.8 ± 2.2	77.1 ± 2.8	15.8 ± 2.2	77.1 ± 2.8	15.5 ± 2.1	75.6 ± 3.0	16.2 ± 2.0	75.9 ± 3.0	16.0 ± 2.4
		1	u			76.9 ± 2.9	17.6 ± 2.3	75.7 ± 3.1	18.9 ± 3.2	75.7 ± 3.2	19.0 ± 3.3	76.8 ± 4.0	22.4 ± 4.3	78.1 ± 4.3	22.9 ± 4.9
		2	u			70.0 ± 3.3	19.8 ± 2.2	72.8 ± 3.4	20.8 ± 3.1	72.9 ± 3.4	20.7 ± 3.2	73.1 ± 5.0	22.2 ± 3.7	73.7 ± 5.5	21.9 ± 4.1
		3	u			67.8 ± 3.4	20.1 ± 2.9	70.9 ± 2.9	20.9 ± 3.0	71.0 ± 3.0	21.1 ± 3.1	71.3 ± 5.2	21.3 ± 3.6	71.1 ± 5.8	20.9 ± 3.9
		3	d			71.2 ± 4.9	16.4 ± 3.1	70.0 ± 5.1	16.8 ± 3.2	70.0 ± 5.1	16.7 ± 3.2	64.3 ± 4.0	15.9 ± 2.0	64.3 ± 3.8	15.9 ± 1.9
		2	d			67.6 ± 2.8	13.2 ± 1.4	65.6 ± 2.9	12.3 ± 1.1	65.7 ± 2.9	12.4 ± 1.0	61.3 ± 3.3	11.6 ± 1.2	61.4 ± 3.4	10.9 ± 1.6
SIFT		1000	0.04	20	1.6	64.1 ± 4.6	15.3 ± 3.6	63.8 ± 4.1	15.5 ± 4.0	63.4 ± 3.9	15.1 ± 3.8	59.0 ± 6.7	13.4 ± 4.0	58.2 ± 6.4	13.1 ± 4.0
		0	0.04	20	1.6	64.1 ± 4.6	15.3 ± 3.6	63.8 ± 4.1	15.5 ± 4.0	63.4 ± 3.9	15.1 ± 3.8	59.0 ± 6.7	13.4 ± 4.0	58.2 ± 6.4	13.1 ± 4.0
		1000	0.04	20	1.2	61.2 ± 5.6	14.7 ± 3.9	61.1 ± 5.0	14.9 ± 4.2	60.8 ± 5.1	14.6 ± 4.0	56.7 ± 6.6	12.3 ± 4.1	56.6 ± 6.2	12.5 ± 3.8
		0	0.04	20	1.2	61.2 ± 5.6	14.7 ± 3.9	61.1 ± 5.0	14.9 ± 4.2	60.8 ± 5.1	14.6 ± 4.0	56.7 ± 6.6	12.3 ± 4.1	56.6 ± 6.2	12.5 ± 3.8
		1000	0.04	10	1.6	53.2 ± 5.8	12.3 ± 3.8	59.2 ± 5.3	14.4 ± 3.2	59.1 ± 5.2	14.5 ± 3.3	56.1 ± 7.5	13.9 ± 4.2	55.5 ± 7.1	13.3 ± 4.1
		0	0.04	10	1.6	53.2 ± 5.8	12.3 ± 3.8	59.2 ± 5.3	14.4 ± 3.2	59.1 ± 5.2	14.5 ± 3.3	56.1 ± 7.5	13.9 ± 4.2	55.5 ± 7.1	13.3 ± 4.1
		1000	0.01	20	1.6	60.1 ± 7.3	17.2 ± 5.3	59.1 ± 7.3	16.8 ± 5.5	58.9 ± 7.4	16.8 ± 5.2	58.8 ± 7.0	15.5 ± 4.6	59.4 ± 6.9	15.6 ± 4.7
		0	0.01	20	1.6	60.1 ± 7.3	17.1 ± 5.3	59.2 ± 7.3	16.6 ± 5.3	59.0 ± 7.4	16.6 ± 5.1	59.2 ± 7.0	15.5 ± 4.6	60.1 ± 6.9	16.1 ± 5.0
		1000	0.01	20	1.2	63.7 ± 7.2	17.9 ± 5.2	60.1 ± 7.3	16.8 ± 4.8	60.1 ± 7.3	16.7 ± 4.8	61.4 ± 6.7	15.4 ± 4.1	62.1 ± 6.4	15.8 ± 3.8
		0	0.01	20	1.2	63.8 ± 7.1	17.8 ± 5.2	60.7 ± 7.2	16.8 ± 4.7	60.6 ± 7.2	16.7 ± 4.7	62.7 ± 6.7	16.5 ± 4.9	63.8 ± 6.5	17.4 ± 4.9
		1000	0.01	10	1.2	66.6 ± 6.0	18.6 ± 5.3	58.7 ± 7.0	16.8 ± 5.2	58.5 ± 6.9	16.3 ± 5.1	58.4 ± 7.2	15.1 ± 4.8	58.5 ± 7.2	15.5 ± 4.8
		0	0.01	10	1.2	66.6 ± 6.0	18.5 ± 5.3	58.7 ± 7.0	16.8 ± 5.1	58.5 ± 6.9	16.3 ± 5.0	58.5 ± 7.2	15.0 ± 4.7	58.8 ± 7.2	16.0 ± 5.0
		1000	0.04	10	1.2	55.0 ± 5.8	13.2 ± 3.4	57.4 ± 5.5	13.6 ± 3.6	57.2 ± 5.7	13.4 ± 3.4	54.6 ± 7.2	12.5 ± 3.7	54.4 ± 6.8	12.6 ± 4.1
		0	0.04	10	1.2	55.0 ± 5.8	13.2 ± 3.4	57.4 ± 5.5	13.6 ± 3.6	57.2 ± 5.7	13.4 ± 3.4	54.6 ± 7.2	12.5 ± 3.7	54.4 ± 6.8	12.6 ± 4.1
		1000	0.01	10	1.6	56.7 ± 8.1	17.1 ± 5.5	56.1 ± 7.3	16.7 ± 5.2	56.0 ± 7.4	16.5 ± 5.1	55.6 ± 7.3	15.4 ± 4.8	56.1 ± 7.4	16.2 ± 5.5
		1000	0.01	10	1.6	56.7 ± 8.1	17.1 ± 5.5	56.1 ± 7.3	16.7 ± 5.2	56.0 ± 7.4	16.5 ± 5.1	55.6 ± 7.3	15.4 ± 4.8	56.1 ± 7.4	16.2 ± 5.5
NN						85.8 ± 5.1	39.0 ± 11.8	88.0 ± 3.1	48.3 ± 9.2	<i>88.0 ± 3.6</i>	<i>50.5 ± 10.5</i>	87.9 ± 3.5	49.1 ± 8.9	87.4 ± 2.9	50.2 ± 9.1

Table S5. Validation AUC and TPR@5%FPR for non-registered and registered data and LBP, HOG, and SIFT descriptors, and our custom CNN using single-channel data and 10-fold cross-validation. Results are reported as mean ± standard deviation across folds. Best AUC in bold, best TPR@5%FPR in italic bold. Parameters: LBP: θ_1 = radius, θ_2 = method (u = uniform, d = default); HOG: θ_1 = pixels per cell, θ_2 = number of orientation bins, θ_3 = cells per block; SIFT: θ_1 = number of features, θ_2 = contrast threshold, θ_3 = edge threshold, θ_4 = sigma.

- **Basic/default mapping.** The full P -bit code obtained by thresholding the P neighbours against the centre pixel is retained unchanged, producing a histogram with 2^P bins [18]. This variant preserves maximum detail but yields a high-dimensional, potentially sparse representation.
- **Uniform mapping.** Codes that contain more than two circular bit transitions $0 \leftrightarrow 1$ are merged into a single “miscellaneous” class, whereas the 58 uniform patterns (for $P=8$) remain distinct. The resulting 59-bin histogram reduces dimensionality by a factor of approximately seven

and is empirically more robust to noise and small rotations.

For HOG, the feature embedding is a flattened vector formed by concatenating contrast-normalized histograms of gradient orientations computed across the image. We systematically vary three hyperparameters:

- **Cell size.** Each cell is a square patch of either 8×8 or 16×16 pixels in the input image.
- **Number of orientation bins.** The histogram can contain 6, 9, or 12 equally spaced bins.

			$c_{\cos}(\bar{\mathbf{u}}_c, \bar{\mathbf{s}}_c)$						$\text{FCN}(\mathbf{u}, \mathbf{s})_{\cos}^+$				
			no registration		$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$		$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$		
κ	w	C	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	
High-resolution time-of-flight mass spectrometer Pegasus® BT 4D													
custom CNN	G	I	3	89.2±3.6	57.2±8.0	90.8±3.4	57.1±9.6	91.0±2.8	58.7±8.4	80.6±5.1	37.8±9.4	80.9±5.3	35.8±9.0
custom CNN	G	I	5	88.5±3.5	54.6±6.3	86.6±4.3	51.9±6.3	87.4±4.1	52.0±6.8	77.7±4.3	34.1±7.7	78.3±4.7	33.3±10.1
custom CNN	G	I	10	89.3±3.9	54.9±9.0	89.4±2.2	55.4±6.2	89.6±3.1	55.9±5.8	79.0±4.3	35.4±14.0	80.5±5.0	37.6±10.4
custom CNN	G	I	20	92.3±2.7	65.4±8.3	91.7±2.6	62.7±8.5	92.1±2.4	62.1±7.6	84.9±5.0	44.3±11.4	83.1±4.9	41.8±11.2
custom CNN	Δ	I	3	90.4±2.5	56.8±10.5	89.0±3.3	53.5±11.6	89.0±3.1	53.7±8.6	82.2±4.4	38.6±10.9	82.1±4.4	37.8±8.8
custom CNN	Δ	I	5	90.0±4.6	61.0±12.3	90.4±3.0	58.8±8.5	89.6±3.4	57.9±10.5	82.1±4.2	40.3±5.0	82.5±4.5	40.1±5.8
custom CNN	Δ	I	10	92.3±2.8	62.5±11.4	89.7±4.6	59.6±8.9	89.9±4.8	58.8±10.2	83.3±4.6	43.3±8.6	83.0±4.9	40.2±10.8
custom CNN	Δ	I	20	91.0±2.6	59.9±6.5	91.3±2.4	58.2±6.5	91.2±2.6	59.8±8.4	81.9±4.1	38.6±9.7	81.6±4.5	37.4±9.6
custom CNN	—	—	1	90.8±3.7	59.6±8.3	91.2±2.7	59.6±8.5	91.2±2.9	59.7±7.3	82.7±4.8	39.6±9.3	82.8±5.2	40.9±7.9
Low-resolution time-of-flight mass spectrometer Pegasus® 4D-C ($\Delta = \mu_{\text{low}} - \mu_{\text{high}}$; σ omitted)													
custom CNN	G	I	3	-10.9	-22.4	-12.4	-22.9	-11.5	-22.9	-4.8	-7.4	-3.8	-4.7
custom CNN	G	I	5	-10.5	-21.6	-9.9	-19.7	-10.9	-19.8	-3.9	-5.6	-4.4	-4.0
custom CNN	G	I	10	-10.4	-17.7	-10.9	-19.6	-11.4	-20.8	-1.8	-1.9	-2.9	-4.4
custom CNN	G	I	20	-12.5	-27.8	-12.2	-26.0	-11.1	-24.1	-6.9	-10.9	-4.7	-7.9
custom CNN	Δ	I	3	-29.0	-35.1	-27.4	-35.0	-27.0	-36.2	-20.7	-23.2	-21.2	-21.6
custom CNN	Δ	I	5	-15.3	-30.8	-17.4	-29.8	-16.5	-28.9	-10.6	-14.2	-10.9	-12.3
custom CNN	Δ	I	10	-14.7	-28.3	-13.3	-27.9	-12.5	-25.8	-9.5	-14.7	-8.0	-11.5
custom CNN	Δ	I	20	-12.7	-24.8	-12.0	-22.1	-12.1	-23.3	-5.4	-6.7	-3.4	-4.4
custom CNN	—	—	1	-11.9	-22.7	-12.7	-23.4	-12.6	-22.3	-6.1	-5.7	-5.6	-6.9

Table S6. Comparison of high- and low-resolution mass spectrometry detectors. Area Under ROC Curve and TPR@5%FPR results for five spatial-alignment strategies on channel representation (Sec. 3.2) of 2D Gas Chromatography Mass Spectrometry measurements on Human Scent Dataset (Sec. 4). Results reported as $\mu \pm \sigma$ across 10 folds for *high-resolution*; *low-resolution* rows show $\Delta = \mu_{\text{low}} - \mu_{\text{high}}$ (negative, red; σ omitted). Identity verification performed with a custom-designed Convolutional Neural Network (CNN). Registration performed with weighted cosine $c_{\cos}(\bar{\mathbf{u}}_c, \bar{\mathbf{s}}_c)$ and fully convolutional-based $\text{FCN}(\mathbf{u}, \mathbf{s})_{\cos}^+$ compound peak detectors introduced in [11] with 22 and 24 detected compounds. Gaussian κ_G and triangular κ_Δ channel representation kernels were used with inversed-intensity I weightings for representation learning, with the number of channel representation kernels $C \in \{1, 3, 5, 10, 20\}$. In addition, a single-channel representation obtained by summing intensities across all m/z values was used as $C = 1$.

- **Block geometry** c . For contrast normalisation, neighbouring cells are grouped into rectangular blocks of either 2×2 or 3×3 cells.

For SIFT, we vary four key hyperparameters while maintaining the default number of layers per octave (3). The final feature vector is constructed by computing the mean of all local 128-dimensional SIFT descriptors, resulting in a fixed-length global representation. Specifically, we explore the following hyperparameters:

- **Number of features** $n \in \{0, 1000\}$. This parameter specifies the maximum number of keypoints retained after detection; $n=0$ disables the cap, whereas $n=1000$ limits the descriptor set to the 1000 strongest responses.
- **Contrast threshold** $f \in \{0.001, 0.04\}$. Filters out weak features in semi-uniform regions.
- **Edge threshold** $e \in \{10, 20\}$. Filters out edge-like features.

S1.8. Comparison of High- and Low-resolution Mass Spectrometers

We evaluate our custom CNN separately on the high- and low-resolution subsets of HSD (Sec. 4) (see Tab S7 for subsets statistics). For each detector, identity- and gender-aware folds are used, containing only samples from that detector. The cohorts are *not identity-paired* (only 3 identities overlap). High-resolution results are reported as $\mu \pm \sigma$ over 10 folds; low-resolution rows report $\Delta = \mu_{\text{low}} - \mu_{\text{high}}$ (with σ omitted). See Table S6.

Discussion. On Pegasus® BT 4D (high-res), the best configuration achieves **92.3 ± 2.7** AUROC and **65.4 ± 8.3** TPR5%FPR with $G, I, C=20$; $\Delta, I, C=10$ is comparable (92.3 ± 2.8/62.5 ± 11.4). Increasing C helps up to 10–20; a single summed channel ($C=1$) is already strong (91.2 ± 2.9/59.7 ± 7.3), so multi-channel yields modest gains. Registration via $\mathcal{R}^{(22)}$ or $\mathcal{R}^{(24)}$ does not improve over no reg-

istration (differences within σ), and the FCN-cosine variant is consistently lower ($\approx 8-10$ AUROC and $\approx 20-24$ TPR points behind across C). Relative to these high-res means, Pegasus® 4D-C (low-res) shows uniformly negative deltas; for $G, I, C=20$ the drop is -12.5 AUROC / -27.8 TPR, with the largest declines at small C and Δ kernels (e.g., $C=3$: $-29.0 / -35.1$).

Overall, high-resolution appears more suitable for custom CNN; however, because cohorts are not matched, a *within-subject paired acquisition* (matched-pairs) study — measuring the same identities on both instruments under identical protocol and timing — is required to attribute differences to detector resolution rather than cohort/batch effects.

Detector resolution	Total num.	Number of identities n	Number of samples per identity			
			min	max	avg	median
low	1341	116	1	76	11.6	10
high	1187	139	2	34	8.5	9

Table S7. Statistics of detector-specific subsets of Human Scent Dataset. Pegasus® 4D-C is the low-resolution detector; Pegasus® BT 4D is the high-resolution detector. Columns: dataset name, number of samples, number of identities n , and per-identity counts (min, max, mean, median). See full dataset statistics in Supplementary Table S1.

S1.9. Regularization and Data Augmentation

To improve generalization under limited data, we regularized the custom CNN with dropout. In the *Backbone*, dropout (0.05) was applied in the second and third stages. In the *Down-sampling Module*, dropout rates of 0.1 and 0.15 were used in the early and late convolutional blocks, respectively (see Sec. S1.5).

We applied random Gaussian noise to the input data during training across all models for identity verification. The noise was added independently to each sample by scaling normally distributed random values with the corresponding percentage of intensities. Exactly, for each encoded measurement \mathbf{C} , we generated noise $\epsilon \sim \mathcal{N}(0, 1)$ of the same shape and scaled it element-wise as $\epsilon \cdot (\gamma_g \cdot \mathbf{C})$ where γ_g controls the relative noise level. In all our experiments, we used $\gamma_g = 0.1$.

We employed channel-wise brightness augmentation during training. For each encoded measurement \mathbf{C} , the brightness in every channel k was randomly scaled by a factor $\mu_{b,k} \sim U(1-\gamma_b, 1+\gamma_b)$ where γ_b controls the maximum relative change. In all our experiments, we fixed $\gamma_b = 0.1$.

We further applied channel-wise contrast augmentation during training. For each encoded measurement \mathbf{C} , the mean intensity of each channel $\overline{\mathbf{C}}_k$ was first computed as $\overline{\mathbf{C}}_k = \frac{1}{|\Omega|} \sum_{\mathbf{u} \in \Omega} \mathbf{C}_{k,\mathbf{u}}$, where Ω denotes the set of all pix-

els in measurement \mathbf{C} . The contrast in channel k was then randomly scaled by a factor $\mu_c \sim U(1-\gamma_c, 1+\gamma_c)$ and applied as $\mathbf{C}'_{k,\mathbf{u}} = (\mathbf{C}_{k,\mathbf{u}} - \overline{\mathbf{C}}_k)\mu_c + \overline{\mathbf{C}}_k$ where γ_c controls the strength. In all our experiments $\gamma_c = 0.1$.

S1.10. Complete Results of CNN Embedding Encoders

In this section, complete results of experiments with CNN embedding encoders are presented in Table S8.

κ	w	C	$c_{\cos}(\bar{\mathbf{u}}_c, \bar{\mathbf{s}}_c)$								$\text{FCN}(\mathbf{u}, \mathbf{s})_{\cos}^+$			
			no registration		$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$		$\mathcal{R}^{(22)}$		$\mathcal{R}^{(24)}$			
			AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%	AUROC	TPR@5%		
custom CNN	G	U	3	88.3±3.8	52.4±9.3	87.2±4.2	47.1±8.0	87.3±3.5	48.6±7.1	84.6±4.5	42.9±6.8	83.4±4.5	40.5±7.8	
custom CNN	G	U	5	89.0±2.6	50.4±8.7	88.0±3.5	50.7±7.9	87.2±3.2	50.2±6.7	85.8±3.4	46.5±5.5	84.4±3.7	41.1±6.2	
custom CNN	G	U	10	88.7±3.9	53.7±9.7	86.2±4.2	46.0±10.0	85.8±4.2	43.6±9.2	82.2±5.2	38.2±8.6	83.2±5.1	39.2±8.5	
custom CNN	G	U	20	89.0±3.7	53.1±7.1	87.5±3.9	49.0±6.6	87.9±3.0	50.3±6.6	84.9±3.9	43.6±5.6	84.2±4.8	43.4±7.1	
GoogLeNet	G	U	3	87.1±3.2	43.9±6.9	86.5±4.0	44.4±7.7	86.6±3.7	43.8±7.7	84.1±4.4	34.9±8.6	84.8±4.3	35.6±9.4	
Inception-v3	G	U	3	86.0±3.5	40.2±5.6	85.5±3.3	39.2±5.0	85.5±3.5	40.2±5.1	83.5±3.9	32.8±6.8	84.0±4.0	34.3±9.1	
ResNet-101	G	U	3	85.5±3.2	39.1±8.2	85.4±3.3	37.0±4.6	85.5±3.5	37.5±6.4	84.3±4.0	33.1±6.6	84.7±4.4	33.5±8.7	
ResNet-18	G	U	3	87.4±3.2	43.3±6.5	86.6±3.1	40.6±5.9	86.9±3.0	41.1±5.9	84.7±4.0	33.6±7.3	85.2±4.1	33.5±7.7	
VGG16	G	U	3	75.0±7.6	46.6±12.2	78.3±9.8	41.9±9.4	78.4±9.9	42.2±8.9	73.8±7.4	27.6±11.2	73.8±7.7	26.6±8.2	
custom CNN	G	I	3	84.2±7.8	44.6±13.7	88.2±2.9	50.8±5.1	86.7±3.9	46.8±8.0	84.0±4.5	40.5±5.6	83.8±5.1	40.0±7.8	
custom CNN	G	I	5	89.5±3.4	53.3±8.3	88.3±3.4	51.3±6.5	88.4±3.4	50.7±6.7	85.6±4.4	43.5±7.5	85.4±3.8	43.9±6.2	
custom CNN	G	I	10	83.6±5.2	38.9±12.2	87.0±4.1	43.2±8.4	84.5±5.2	41.6±8.7	83.0±4.7	36.8±9.8	83.3±4.9	39.5±8.3	
custom CNN	G	I	20	84.7±4.1	48.1±3.4	87.1±2.9	49.5±5.8	86.1±4.7	48.1±5.5	88.5±1.1	53.6±2.3	87.0±4.1	46.2±1.7	
GoogLeNet	G	I	3	87.9±3.5	46.8±7.7	86.8±3.9	44.3±8.4	86.7±3.4	44.7±7.5	85.0±3.8	35.2±8.9	84.9±4.1	36.1±10.3	
Inception-v3	G	I	3	85.4±3.5	43.0±9.0	84.3±3.6	39.8±5.6	84.5±3.7	40.6±6.2	82.6±3.9	32.4±9.1	82.8±3.7	33.6±10.8	
ResNet-101	G	I	3	86.3±4.1	41.5±11.4	86.0±4.2	38.5±8.7	86.5±4.1	40.6±7.7	84.5±4.7	35.3±8.7	84.9±4.4	35.2±9.1	
ResNet-18	G	I	3	87.5±3.9	44.9±8.8	87.3±3.5	42.9±5.2	86.9±3.2	41.9±5.9	85.2±3.8	36.1±7.1	85.5±4.3	36.2±8.2	
VGG16	G	I	3	76.5±7.3	47.3±12.9	79.5±9.6	42.9±10.4	79.8±9.4	42.2±9.4	74.6±7.2	26.2±8.3	74.5±6.9	26.6±6.6	
custom CNN	Δ	U	3	55.4±4.9	3.4±4.2	55.6±5.3	3.3±4.2	55.7±4.5	3.8±3.6	55.5±5.0	3.3±3.5	55.6±4.8	3.5±3.9	
custom CNN	Δ	U	5	55.5±4.6	3.4±4.3	55.6±5.5	3.5±4.3	55.4±4.7	3.4±3.8	54.9±4.4	3.6±3.7	55.0±3.8	3.3±3.8	
custom CNN	Δ	U	10	55.4±4.7	3.7±4.1	55.3±4.9	3.6±4.2	54.8±4.3	3.2±3.9	55.4±4.2	3.9±4.5	55.7±5.1	3.2±3.7	
custom CNN	Δ	U	20	55.8±5.7	3.6±4.5	55.1±4.2	3.7±3.8	55.7±5.1	3.8±4.7	54.8±4.0	3.4±4.2	55.0±4.2	3.4±3.8	
GoogLeNet	Δ	U	3	53.0±3.0	1.2±1.7	52.9±2.7	1.2±1.4	53.2±3.3	1.7±2.0	52.9±3.0	1.3±1.6	53.0±3.1	1.3±1.6	
Inception-v3	Δ	U	3	56.0±5.4	3.4±4.2	55.9±5.3	3.3±4.2	56.0±5.1	3.6±3.8	55.6±5.0	3.2±3.6	55.9±5.3	3.4±4.0	
ResNet-101	Δ	U	3	52.8±2.5	1.2±1.7	52.6±2.3	1.2±1.4	52.6±2.5	1.7±2.0	52.4±2.1	1.3±1.6	56.1±11.1	5.0±11.5	
ResNet-18	Δ	U	3	51.9±1.9	0.7±0.8	51.9±1.8	0.8±0.9	52.1±2.4	1.2±1.8	51.9±1.9	0.9±0.9	52.1±2.4	0.9±1.2	
VGG16	Δ	U	3	52.4±2.3	1.3±1.7	52.5±2.3	1.2±1.4	52.6±2.9	1.9±1.9	52.4±2.4	1.5±1.5	52.7±2.8	1.4±1.5	
custom CNN	Δ	I	3	83.1±8.4	42.9±13.7	81.0±5.1	35.2±6.8	80.4±4.7	36.2±5.0	78.7±5.9	31.8±6.6	77.8±6.1	29.8±7.6	
custom CNN	Δ	I	5	82.3±7.0	43.5±5.1	85.5±4.1	43.5±7.4	85.4±3.6	43.9±6.8	80.6±5.7	34.6±9.2	77.8±6.3	27.6±9.7	
custom CNN	Δ	I	10	87.1±1.3	51.3±4.8	86.8±3.7	52.4±5.9	86.3±3.9	50.0±6.8	82.7±4.9	42.2±8.4	81.4±5.0	37.0±9.6	
custom CNN	Δ	I	20	82.3±5.2	37.4±10.5	81.5±3.1	51.3±4.2	87.3±2.5	48.1±5.1	88.1±1.1	51.2±3.5	86.4±2.4	49.1±3.7	
GoogLeNet	Δ	I	3	84.0±3.3	36.4±8.0	84.1±3.8	32.2±6.1	83.9±3.6	31.5±6.0	82.0±3.2	25.3±7.1	81.7±3.1	23.3±5.3	
Inception-v3	Δ	I	3	83.4±3.4	35.7±8.4	82.5±3.8	28.7±10.3	83.0±3.5	28.4±8.5	79.6±3.6	24.1±9.1	79.7±3.3	22.7±8.6	
ResNet-101	Δ	I	3	86.0±3.4	40.4±9.7	83.6±2.9	32.5±4.8	83.5±3.3	32.6±6.0	81.7±1.8	27.2±4.1	81.4±1.9	27.7±4.1	
ResNet-18	Δ	I	3	83.5±4.1	33.3±9.8	83.0±4.3	26.2±9.4	83.0±3.9	26.4±9.5	80.6±3.9	21.4±7.6	81.0±3.9	21.2±7.9	
VGG16	Δ	I	3	77.0±4.6	40.5±6.6	77.4±3.4	38.1±5.9	77.4±3.4	38.0±6.0	71.7±2.7	25.1±3.4	70.8±2.8	25.2±3.6	

Table S8. All cross-validation Area Under ROC Curve and TPR@5%FPR results for five spatial-alignment strategies on channel representation (Sec. 3.2) of 2D Gas Chromatography Mass Spectrometry measurements on Human Scent Dataset (Sec. 4). Results reported as $\mu \pm \sigma$ across 10 folds. Identity verification performed with: (i) custom-designed Convolutional Neural Network (CNN), and (ii) five pre-trained neural network backbones *ResNet-18* [10], *ResNet-101* [10], *VGG-16* [32], *Inception-v3* [34], and *GoogLeNet* [34] with appended and trained projection layer. Registration performed with weighted cosine $c_{\cos}(\bar{\mathbf{u}}_c, \bar{\mathbf{s}}_c)$ and fully convolutional-based $\text{FCN}(\mathbf{u}, \mathbf{s})_{\cos}^+$ compound peak detectors introduced in [11] with 22 and 24 detected compounds. Gaussian κ_G and triangular κ_{Δ} channel representation kernels were used with uniform U and inversed-intensity I weightings for representation learning, number of channel representation kernels $C \in \{1, 5, 10, 20\}$.