Advances in Science, Technology and Engineering Systems Journal
Vol. 11, No. 4, 1-34 (2026)
www.astesj.com
Special Issue on Emerging Multidisciplinary Directions in Engineering, Computing, and Applied Sciences

ASTES Journal
ISSN: 2415-6698

Payload and Watermarking Trade-Offs in Online Handwriting: A Structured Comparative Study

Ruben Copado-Torreblanca, Marcos Faundez-Zanuy*

Technology Department, Tecnocampus, Universitat Pompeu Fabra, Mataró (Barcelona), Spain

Email(s): rcopadoo@tecnocampus.cat (R. Copado-Torreblanca), faundez@tecnocampus.cat (M. Faundez-Zanuy)

*Corresponding Author: Marcos Faundez-Zanuy, Technology Department, Tecnocampus, Universitat Pompeu Fabra, Mataró (Barcelona), Spain. Email: faundez@tecnocampus.cat

ARTICLE INFO

Article history:
Received: 02 May, 2026
Revised: 23 July, 2026
Accepted: 25 July, 2026
Online: 31 July, 2026

Keywords:
Online handwriting
Digital watermarking
Payload capacity
Signal distortion
Robustness analysis
Biometric security

ABSTRACT

Online handwriting is a dynamic signal that can support digital watermarking, but existing studies have mainly evaluated individual embedding methods, leaving limited structured evidence on the joint relationship between payload, distortion, robustness, and host-signal characteristics. This paper addresses this gap through a comparative framework that combines cross-dataset host-signal characterization with a controlled evaluation of representative watermarking families. A heterogeneous set of handwriting and signature datasets is characterized in terms of trajectory length, task type, and in-air/on-surface composition. LSB, Difference Expansion, QIM, DCT, DWT, DFT, and SVD are then examined under a common reference setting. Payload is analyzed at method, dataset, and task levels, whereas geometric distortion and robustness are experimentally evaluated on a reference online signature under additive Gaussian noise and low-pass filtering. Within the evaluated framework, the controlled experiment reveals method-dependent payload–distortion–robustness trade-offs, while the cross-dataset analysis shows substantial variation in nominal host-signal support across trajectory structures.

1. Introduction

Online handwriting is a dynamic signal that captures temporal and kinematic information generated during the writing process, such as trajectory, velocity, acceleration, and pressure. Due to its compatibility with digital acquisition devices and its relevance in biometric, forensic, and signal protection scenarios, online handwriting has been widely studied in applications related to authentication, human behavior analysis, and digital document processing.

One of the earliest and most influential state-of-the-art works in this area is the review by Faundez-Zanuy [1], which provided a comprehensive overview of online signature recognition systems at a time when statistical modeling and template-based matching techniques dominated the literature. That work described the main system components, including feature extraction, matching strategies such as Dynamic Time Warping (DTW) and Hidden Markov Models (HMMs), and the role of evaluation protocols and error metrics, while also emphasizing the challenge posed by skilled forgeries and the behavioral nature of handwriting as a biometric modality.

However, the current context differs from that earlier stage in an important way. Beyond recognition itself, there is an increasing interest in protecting online handwriting signals as digital objects that may carry sensitive or valuable information. This has motivated the use of watermarking techniques capable of embedding data into the signal while attempting to preserve its utility and limit the induced distortion.

Existing research on watermarking for dynamic biometric signals has largely focused on individual embedding techniques or on application-specific implementations. Consequently, there is still limited structured evidence showing how different watermarking families compare in terms of payload, geometric distortion, and robustness when applied to online handwriting. Moreover, the influence of host-signal properties—including trajectory length, task type, and in-air/on-surface composition—has received comparatively little attention, despite its potential effect on the amount and location of embeddable information.

This paper does not propose a new watermarking algorithm. Its original contribution lies in organizing representative spatial, reversible, quantization-based, and transform-domain methods within a common comparative framework and relating their behavior to the structural properties of online handwriting signals. To this end, the study combines: (i) a cross-dataset characterization of handwriting and signature signals; (ii) a method-, dataset-, and task-level analysis of nominal payload; and (iii) a controlled evaluation of geometric distortion and robustness on a reference online signature. This design provides structured comparative evidence while explicitly distinguishing cross-dataset signal characterization from reference-level experimental validation.

1.1. Research Questions

In order to address this gap, the present study is guided by the following research questions:

- **RQ1:** How does the nominal payload capacity differ across watermarking methods commonly applicable to online handwriting signals?
- **RQ2:** How does watermark embedding affect the geometric integrity of the host handwriting trajectory, and how does this distortion evolve with embedding strength?
- **RQ3:** Which watermarking methods provide greater robustness against additive noise and low-pass filtering in online handwriting signals?
- **RQ4:** To what extent do host-signal descriptors such as trajectory length, task type, and in-air/on-surface composition suggest differences in the practical watermarking potential of online handwriting data?

1.2. Main Contributions

The novelty of this work is methodological and integrative rather than algorithmic: instead of introducing another isolated watermarking scheme, the study organizes representative embedding families and heterogeneous online handwriting signals within a common comparative perspective. The main contributions are:

- A cross-dataset characterization specifically oriented toward watermarking, covering trajectory length, task category, and in-air/on-surface composition across heterogeneous handwriting and signature corpora.
- A unified comparison of seven representative watermarking approaches—LSB, Difference Expansion, QIM, DCT, DWT, DFT, and SVD—covering spatial, reversible, quantization-based, and transform-domain families.
- A multilevel analysis of nominal payload that distinguishes method-level capacity from the host-signal availability associated with datasets and task categories.
- A controlled reference-level evaluation of geometric distortion and robustness using common metrics and attack conditions, while explicitly delimiting the scope of the experimental evidence.
- A structured interpretation of the payload–distortion–robustness trade-off, including the observation that low distortion does not necessarily imply robustness, that robustness is attack-dependent, and that host-signal descriptors may condition practical embedding opportunities.

1.3. Paper Organization

The remainder of the paper is organized as follows. Section 2 reviews the main background on online handwriting as a dynamic signal and on watermarking methods relevant to biometric and handwriting data. Section 3 presents the datasets considered in the study and analyzes their main structural properties from the perspective of watermark embedding. Section 4 introduces the watermarking methods under comparison, and Section 5 describes the experimental framework. Section 6 provides a descriptive analysis of the host signals. Section 7 examines payload capacity from both method-level and dataset/task-level perspectives. Section 8 analyzes the geometric distortion introduced by watermark embedding, and Section 9 evaluates robustness under Gaussian noise and low-pass filtering. Finally, Sections 10–13 present the discussion, limitations, conclusions, and future research directions.

2. Related Work

This section presents the main background relevant to this study. First, online handwriting is considered as a dynamic signal, since its trajectory, temporal evolution, and pressure information determine both its analytical value and its suitability as a host signal for watermark embedding. The section then reviews watermarking in biometric and handwriting signals, introducing the principal method families commonly discussed in the literature, such as spatial, transform-domain, quantization-based, and reversible approaches, as well as the main design dimensions associated with them, including blind versus non-blind detection and fragile versus robust embedding. This context serves to position the present work and to define the specific gap addressed by the proposed comparative analysis.

2.1. Online Handwriting as a Dynamic Signal

Online handwriting and on-line signature are dynamic signals acquired during the writing process, where the recorded information goes beyond the final visible trace and includes temporal and kinematic variables such as trajectory, velocity, acceleration, and pen pressure. This dynamic nature distinguishes online modalities from offline handwriting and enables the use of time-dependent representations for modeling human writing behavior [1].

A classical description of online signature systems includes the stages of acquisition, preprocessing, feature extraction, matching, and decision-making [1]. Within this framework, feature representations may rely on raw time functions, derivative-based descriptors, or more compact statistical summaries derived from transformed signal domains. Matching strategies range from elastic distance measures, such as Dynamic Time Warping (DTW), to statistical classifiers and learning-based approaches. In biometric applications, system performance is commonly evaluated through False Acceptance Rate (FAR), False Rejection Rate (FRR), and

Detection Error Tradeoff (DET) curves, which provide a standard view of the trade-off between security and usability [2].

Several works have refined the modeling of online signature dynamics. A representative example is the combination of Vector Quantization and DTW, where the signature is transformed into a codeword sequence before temporal alignment, reducing computational cost while preserving time-warping capability [3]. Other studies have enriched the signal representation by incorporating first- and second-order temporal derivatives, improving discrimination under intra-user variability and forgery conditions [4]. Rotation-invariant dynamic descriptors derived from position, velocity, and acceleration have also been proposed to increase robustness against geometric variability [5], while explicit rotation, scaling, and translation normalization has been used as a preprocessing strategy to reduce mismatch due to global distortions [6].

In addition to general verification performance, part of the literature has focused on difficult biometric scenarios such as skilled forgery detection. In this direction, the integration of DTW with Sigma LogNormal-based kinematic descriptors has shown that neuromuscular information can complement conventional time functions and improve the separation between genuine signatures and forgeries [7]. Related work has also examined inclination-related differences between authentic and forged signatures, further supporting the relevance of dynamic and behavioral information in online signature analysis [8].

Taken together, these works show that online handwriting is not only a geometric trace, but a structured temporal signal whose properties make it suitable for recognition, behavioral analysis, and signal protection tasks. This dual nature is especially relevant in the present study, where the host signal is analyzed not from the viewpoint of verification alone, but also from the perspective of its capacity to support embedded information.

2.2. Watermarking in Biometric and Handwriting Signals

Beyond recognition, recent research has increasingly considered the need to protect dynamic biometric signals as digital objects that may contain sensitive or valuable information. In this context, watermarking provides a mechanism for embedding information into the signal while attempting to preserve its usability and control the distortion introduced during embedding.

A broad conceptual foundation for this area is provided by the information hiding literature, where digital watermarking is framed together with related problems such as steganography, fingerprinting, and covert communication. A key idea emphasized in this context is that no watermarking method is universally optimal, since practical designs must balance payload, transparency, robustness, and application-specific requirements [9]. This trade-off is particularly relevant in online handwriting, where the host signal is itself variable, dynamic, and potentially biometrically meaningful.

Within online signature data, watermarking has been explored through several signal-oriented approaches. Transform-domain embedding based on the Discrete Cosine Transform (DCT) has been proposed for online signatures in both single-bit and multi-bit settings [10]. A related blind DCT-based formulation has also been introduced to enable watermark extraction without access to the original signal [11]. In contrast, spatial-domain reversible or near-reversible embedding strategies such as Least Significant Bit (LSB) substitution and Difference Expansion (DE) have been studied specifically in online signature signals, with emphasis on their embedding behavior and signal alteration [12]. More recently, fragile watermarking has also been proposed directly in the pressure channel of online signatures for integrity checking, with a specific embedding rule designed to preserve pen-up samples ($p = 0$) and avoid spurious strokes in the reconstructed trajectory [13]. From a different perspective, template protection has also been addressed through fixed-length, quantized, and non-invertible representations, as illustrated by SigTem, which aims to provide revocability and diversity without directly embedding information into the signal itself [14].

Related developments in biomedical signals help contextualize these approaches. Watermarking in EEG and ECG has been studied under constraints similar to those of online handwriting, namely the need to preserve the practical utility of the host signal while embedding information robustly or reversibly [15, 16, 17, 18, 19]. At the medical imaging level, broader reviews have also emphasized the recurring trade-offs between imperceptibility, robustness, payload, and recoverability [20]. Although these signals differ from online handwriting in structure and application, they reinforce the general idea that host-signal characteristics strongly condition feasible embedding strategies.

Overall, the literature shows that watermarking in biometric and signal-based applications is moving from generic multimedia formulations toward more application-dependent designs. In online handwriting, this shift is especially important because the signal is not only a carrier of information, but also a structured representation of human motor behavior.

2.3. Families of Watermarking Methods

The watermarking literature commonly distinguishes several method families according to the embedding domain and the detector assumptions. Introductory and review works consistently separate spatial-domain techniques from transform-domain approaches, and identify imperceptibility, robustness, payload capacity, security, and computational complexity as the main criteria for method comparison [21, 22, 23, 24].

Spatial-domain methods operate directly on the host samples. In the context of online handwriting, the most representative examples are LSB substitution and Difference Expansion. These approaches are attractive because of their simplicity and, in some cases, reversibility, but they are generally more sensitive to perturbations when compared with robust transform-domain formulations [12]. Their main strength lies in high local controllability of the embedding process, which makes them relevant when exact or near-exact signal recovery is desired.

Transform-domain methods embed information after projecting the host signal into a frequency or multiresolution representation. In the broader watermarking literature, DCT, DWT, DFT, and SVD-based methods are recurrent because they often provide a more favorable robustness–transparency balance than direct sample modification [21, 22, 23]. This general tendency is also reflected in application-specific schemes combining DCT, DFT, SVD, DWT, PCA, or related transforms for image and video watermarking [25, 26, 27, 28, 29, 30]. In online signatures, DCT-based embedding has already shown the practical relevance of transform-domain approaches for dynamic coordinate signals [10, 11].

A further family is formed by quantization-based methods, where information is embedded by forcing the host signal toward quantization regions associated with different watermark symbols. Quantization Index Modulation (QIM) is a representative example of this idea and has been widely used in signal and multimedia watermarking because it offers a clear and controllable robustness–distortion trade-off [28]. From an analytical viewpoint, these methods are especially attractive in comparative studies because their embedding strength is directly governed by interpretable quantization parameters.

Finally, reversible watermarking deserves separate attention because it addresses scenarios in which the exact recovery of the original host signal is important. This is particularly relevant in signals where the embedded data should not permanently alter the source, as illustrated by reversible ECG watermarking and DE-based embedding strategies [19, 12]. In online handwriting, reversible methods are of interest not only because they preserve the signal, but also because they impose specific constraints on achievable payload and robustness.

These families are not merely taxonomic categories; they correspond to different design priorities. Spatial and reversible methods tend to emphasize direct control and recoverability, transform methods tend to emphasize robustness, and quantization-based methods offer an explicit control of the robustness–distortion compromise. This distinction is central to the present work, where the comparison is organized around payload capacity and its interaction with distortion and robustness.

2.4. Research Gap

The reviewed literature provides a solid foundation on both dynamic online handwriting analysis and watermarking design. On the one hand, biometric and signature studies have shown how trajectory dynamics, temporal derivatives, kinematic descriptors, and geometric normalization can improve the analysis of online signatures under variability and forgery conditions [1, 3, 4, 5, 6, 7, 8]. On the other hand, the watermarking literature has established the main design dimensions of embedding methods, including payload, transparency, robustness, and reversibility, across multimedia, biomedical, and signature-oriented applications [9, 21, 22, 23, 24, 20, 10, 11, 12].

However, the studies reviewed above mainly examine individual watermarking methods, specific embedding domains, or application-dependent implementations. As a result, the online handwriting literature still provides limited structured comparative evidence on three connected questions: how payload differs across watermarking families, how increased embedding strength affects geometric distortion and robustness, and how the available host-signal support varies with trajectory length, task type, and in-air/on-surface composition. These dimensions are often discussed separately, which makes it difficult to interpret their combined practical implications.

The present work addresses this gap by integrating two complementary levels of evidence. First, multiple handwriting and signature datasets are characterized to quantify differences in host-signal structure and nominal embedding support. Second, representative watermarking families are evaluated under a controlled reference setting using common distortion and robustness metrics. This combination does not constitute a complete cross-dataset experimental benchmark; rather, it provides a structured baseline for relating method-level trade-offs to the broader characteristics of online handwriting signals.

3. Materials: Datasets and Signal Characterization

This section characterizes the online handwriting and signature datasets considered in the study. The selected corpora differ in acquisition purpose, task design, trajectory length, and in-air/on-surface composition. These properties are relevant because they determine the amount and distribution of host-signal support potentially available for watermark embedding. The section first explains the dataset-selection criteria and then reports the main structural descriptors for each corpus.

3.1. Dataset Selection and Characterization

The datasets were selected according to three criteria: the availability of dynamic coordinate and pressure signals, diversity in trajectory length and task type, and complementarity of application contexts. The selected corpora span signatures, words, extended texts, repetitive patterns, and geometric drawings acquired in biometric-security, health-related, and handwriting-analysis scenarios. The following subsections characterize each dataset in terms of acquisition context, task structure, trajectory length, and in-air/on-surface composition. Pressure samples with $p = 0$ are considered in-air, whereas samples with $p > 0$ are considered on-surface; trajectory length is used as an indicator of the nominal signal support available for watermark embedding.

Each dataset subsection begins with a brief introduction and follows the same structure: sample-length and pressure distributions, shortest and longest recordings, and task- or class-level summaries with representative examples when applicable.

3.1.1. EMOTHAW

EMOTHAW is an online handwriting dataset composed of heterogeneous tasks that combine geometric drawing and handwriting production [31]. From the perspective of the present work, it is relevant because it includes signals with markedly different lengths, pressure usage patterns, and proportions of in-air and on-surface points. This variability makes it suitable for analyzing how watermarking behavior changes as a function of task structure and host-signal characteristics.

EMOTHAW was originally designed to study the relationship between online handwriting and emotional state, particularly anxiety, stress, and depression, assessed using the Depression Anxiety Stress Scales (DASS). This context also explains the diversity of drawing and handwriting tasks included in the dataset.

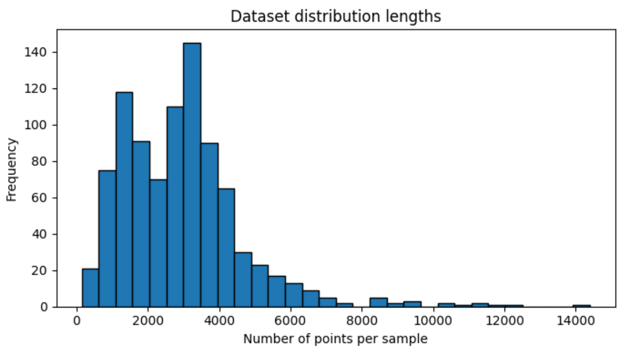


Figure 1: EMOTHAW length distribution (frequency of samples vs. number of points).

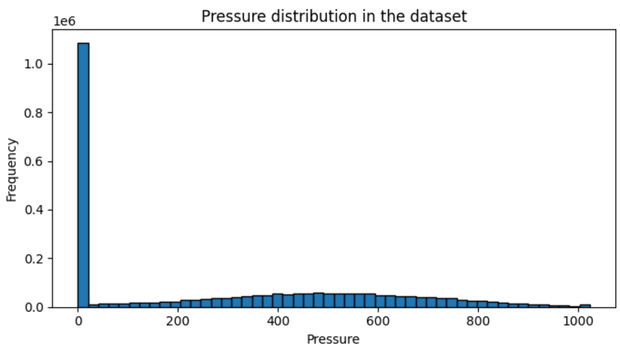
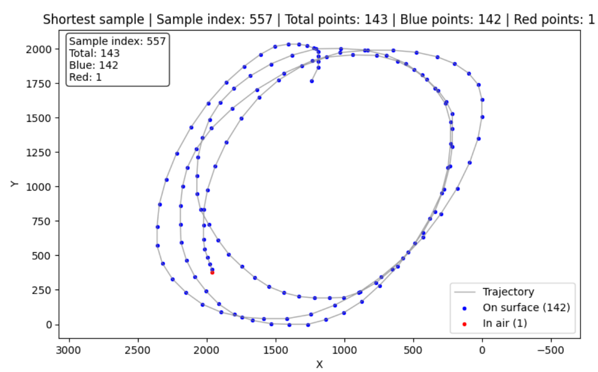
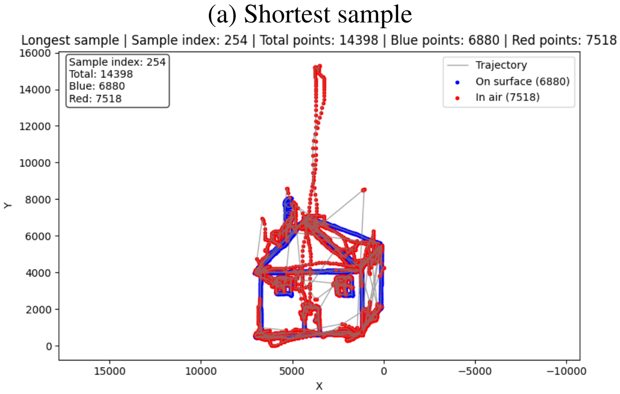


Figure 2: EMOTHAW pressure distribution (frequency of samples vs. pressure range). The dominance of $p = 0$ indicates a large amount of in-air trajectory points.



(a) Shortest sample



(b) Longest sample

Figure 3: EMOTHAW extreme-length samples: (a) shortest and (b) longest trajectories in number of points.

A task-wise summary of sample length is provided in Table 1. The number of samples is highly consistent across tasks, with approximately 128–129 samples per task, while the average trajectory length changes substantially. Tasks such as t2_data and t3_data contain longer average trajectories, whereas t4_data and t5_data are notably shorter. These differences suggest that EMOTHAW covers tasks with distinct temporal and graphomotor demands, which is useful for comparative watermarking analysis.

Table 1: EMOTHAW summary by task: number of samples and length statistics (in points).

Task	Samples	Mean length	Min	Max
t1_data	129	2568.58	755	11841
t2_data	129	4416.14	1813	14398
t3_data	129	3918.14	2150	8465
t4_data	128	1378.93	234	3527
t5_data	129	1202.18	143	2421
t6_data	129	3811.75	1369	11541
t7_data	129	3290.97	1544	8626

Table 2 details the average number of in-air points ($p = 0$) and on-surface points ($p > 0$) per task, together with their relative proportions. This table complements the global pressure histogram in Figure 2 by showing that the balance between in-air and on-surface trajectory segments varies markedly across tasks. In particular, t4_data and t5_data are almost entirely on-surface, with in-air proportions below 2.13%, whereas t6_data contains a majority of in-air points (53.71%). Tasks t1_data, t2_data, and t3_data also exhibit substantial in-air activity, indicating frequent pen lifts or long inter-stroke transitions.

Table 2: EMOTHAW per-task mean number of points in-air ($p = 0$) and on-surface ($p > 0$), including their percentage over total points.

Task	**Samples**	**Mean in-air**	**Mean on-surface**	**Mean total**	**% in-air**	**% on-surface**
t1_data	129	1123.59	1444.99	2568.58	43.74	56.26
t2_data	129	2007.17	2408.97	4416.14	45.45	54.55
t3_data	129	1888.78	2029.36	3918.14	48.21	51.79
t4_data	128	29.43	1349.50	1378.93	2.13	97.87
t5_data	129	9.90	1192.28	1202.18	0.82	99.18
t6_data	129	2047.30	1764.45	3811.75	53.71	46.29
t7_data	129	1261.67	2029.29	3290.97	38.34	61.66

The diversity of task content is illustrated in Figures 4–10, which provide one representative example for each acquisition task. EMOTHAW includes both drawing and handwriting activities: Task 1 corresponds to Folstein's pentagon copying, Task 2 a house, Task 3 four words, Tasks 4 and 5 circle-loop patterns, Task 6 a clock, and Task 7 a full sentence. This range of elicited content is consistent with the variability observed in Tables 1 and 2, and makes EMOTHAW particularly useful for studying how watermarking behavior changes across tasks with different structural and temporal characteristics.

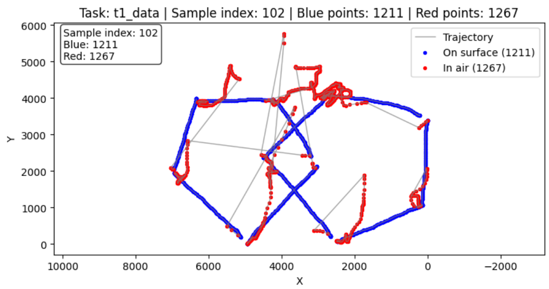


Figure 4: EMOTHAW task example: Task 1, Folstein's pentagon copying.

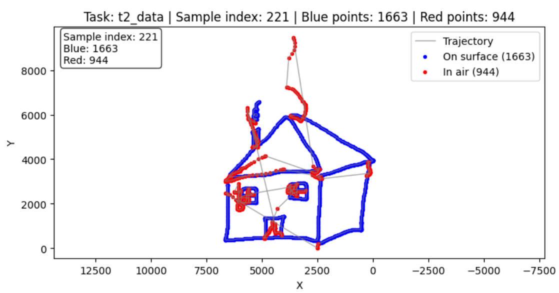


Figure 5: EMOTHAW task example: Task 2, a house.

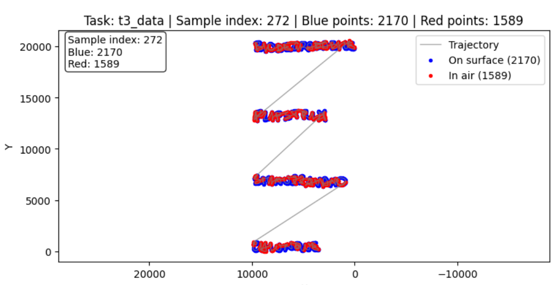


Figure 6: EMOTHAW task example: Task 3, four words.

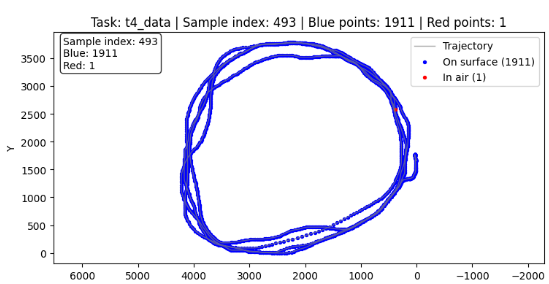


Figure 7: EMOTHAW task example: Task 4, circle loops.

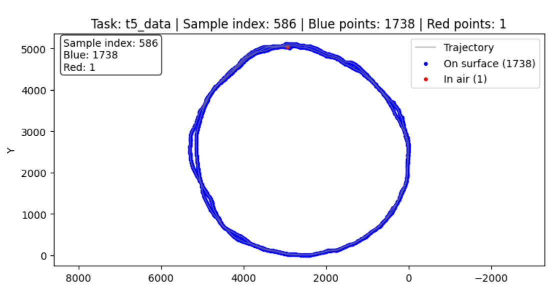


Figure 8: EMOTHAW task example: Task 5, circle loops.

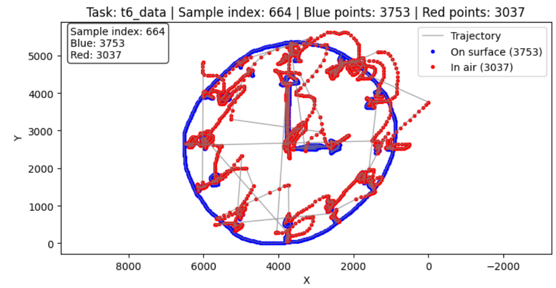


Figure 9: EMOTHAW task example: Task 6, a clock.

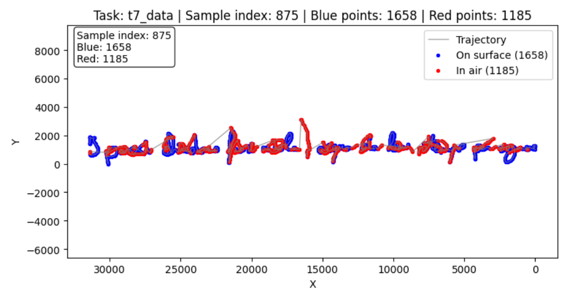


Figure 10: EMOTHAW task example: Task 7, a full sentence.

3.1.2. PaHaW

PaHaW is an online handwriting dataset composed of both drawing and handwriting tasks with different levels of complexity [32]. From the perspective of the present study, it is particularly relevant because it combines relatively simple trajectories with more extended handwriting productions, thereby providing signals with different lengths, pressure usage patterns, and in-air versus on-surface compositions. This variability makes it suitable for analyzing how watermarking performance depends on the structural properties of the host signal.

An additional aspect that makes PaHaW particularly relevant is its original clinical motivation. The dataset was designed to study the relationship between online handwriting and Parkinson's disease by comparing handwriting samples from patients and healthy controls. From the perspective of the present work, this application context is also valuable because it provides a set of tasks with different temporal and structural properties acquired in a neurologically meaningful scenario.

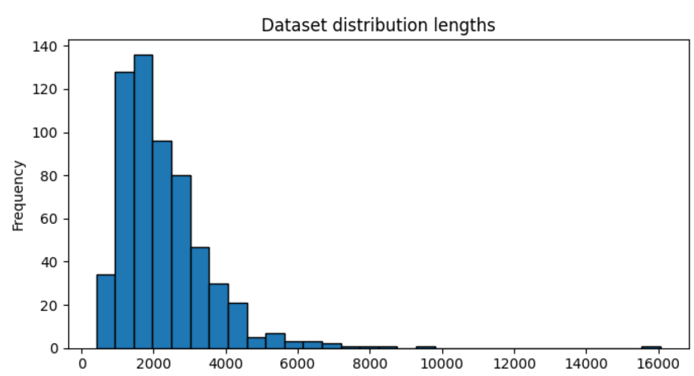


Figure 11: PaHaW length distribution (frequency of samples vs. number of points).

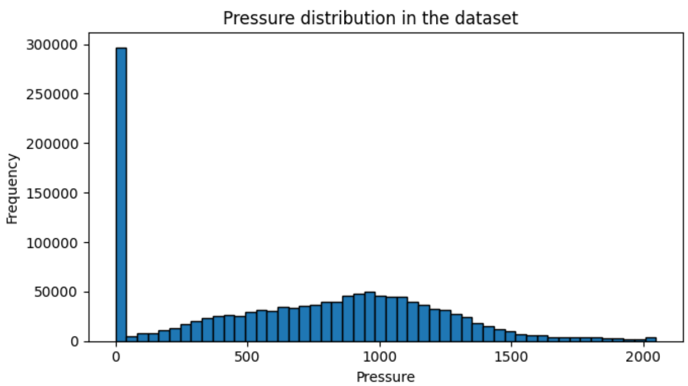
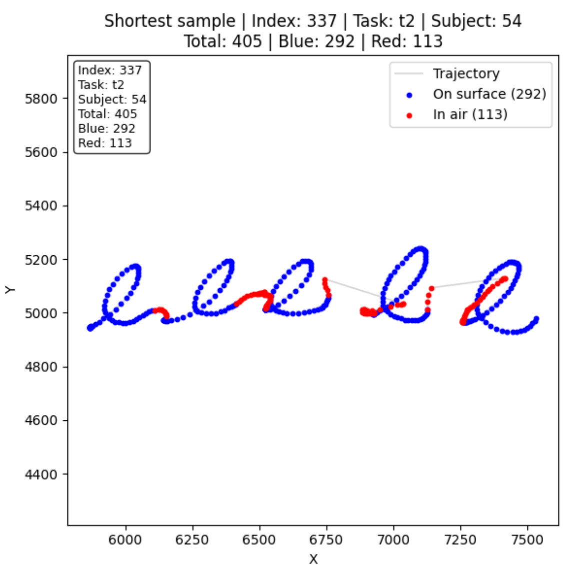
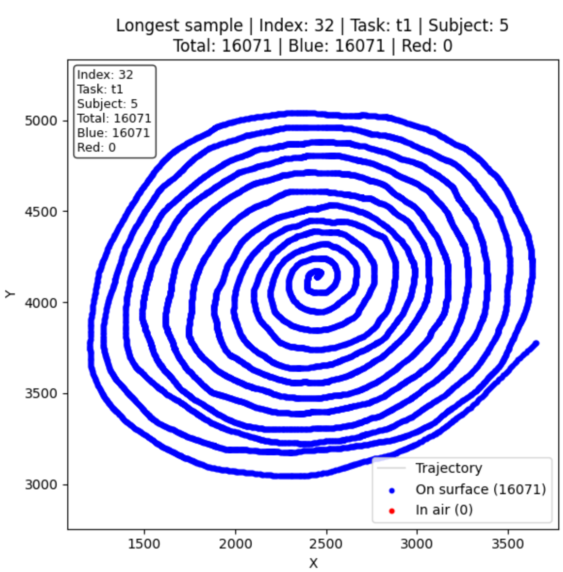


Figure 12: PaHaW pressure distribution (frequency of samples vs. pressure range).



(a) Shortest sample



(b) Longest sample

Figure 13: PaHaW extreme-length samples: (a) shortest and (b) longest trajectories in number of points.

A task-wise summary of trajectory length is provided in Table 3. The number of samples is highly consistent across tasks, with 72 samples for t1 and 75 for the remaining tasks, while the average trajectory length changes substantially depending on the elicited content. Task 1 presents the largest maximum length, reaching 16071 points, whereas the remaining tasks show more moderate maxima. In terms of mean length, t8 and t1 correspond to the longest tasks, while t7 is the shortest. These differences indicate that PaHaW covers writing activities with distinct temporal and structural demands.

Table 3: PaHaW summary by task: number of samples and length statistics (in points).

Task	Samples	Mean length	Min	Max
t1	72	2873.69	552	16071
t2	75	1668.01	405	4226
t3	75	1984.16	693	6615
t4	75	2305.01	614	6827
t5	75	2608.48	855	7993
t6	75	2315.08	796	5783
t7	75	1469.29	585	4423
t8	75	3086.13	1234	7676

Table 4 reports the average number of in-air points ($p = 0$) and on-surface points ($p > 0$) for each task, together with their percentages over the total. This task-level view complements the global pressure histogram in Figure 12 and shows that the proportion of in-air activity depends strongly on the type of task. The spiral task (t1) is almost entirely on-surface, with 97.79% of its points recorded in contact with the surface, whereas the handwriting tasks exhibit a much higher proportion of in-air points due to pen lifts and inter-stroke transitions. Among them, t8 has the highest in-air percentage (34.91%), while t7 has the lowest among the handwriting tasks (17.28%).

Table 4: PaHaW per-task mean number of points in-air ($p = 0$) and on-surface ($p > 0$), including their percentage over total points.

Task	Samples	Mean in-air	Mean on-surface	Mean total	% in-air	% on-surface
t1	72	63.64	2810.06	2873.69	2.21	97.79
t2	75	548.61	1119.40	1668.01	32.89	67.11
t3	75	487.93	1496.23	1984.16	24.59	75.41
t4	75	524.55	1780.47	2305.01	22.76	77.24
t5	75	479.12	2129.36	2608.48	18.37	81.63
t6	75	496.47	1818.61	2315.08	21.44	78.56
t7	75	253.91	1215.39	1469.29	17.28	82.72
t8	75	1077.48	2008.65	3086.13	34.91	65.09

The task inventory is illustrated in Figures 14–21, which provide one representative example for each task. PaHaW includes a geometric drawing task and several handwriting exercises with increasing complexity. Task 1 corresponds to an Archimedean spiral, Tasks 2–4 consist of constrained repetitive cursive patterns, and Tasks 5–8 correspond to word-writing tasks. This progression from highly constrained trajectories to richer handwriting productions is consistent with the variability observed in Tables 3 and 4, and makes PaHaW useful for studying how watermarking behavior changes across tasks with different temporal and structural characteristics.

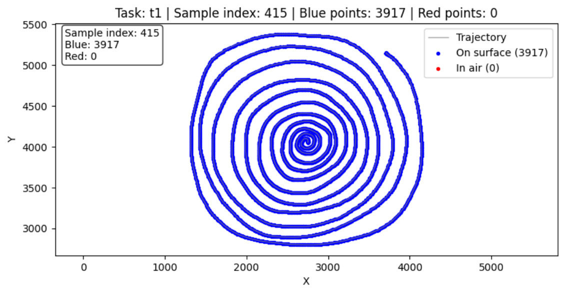


Figure 14: PaHaW task example: Task 1, Archimedean spiral.

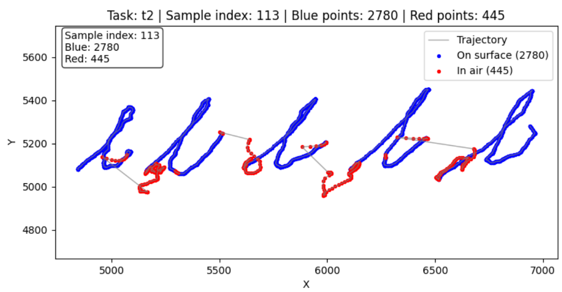


Figure 15: PaHaW task example: Task 2, repetitive cursive pattern.

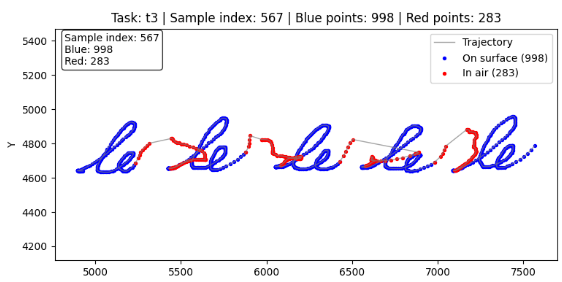


Figure 16: PaHaW task example: Task 3, repetitive cursive pattern.

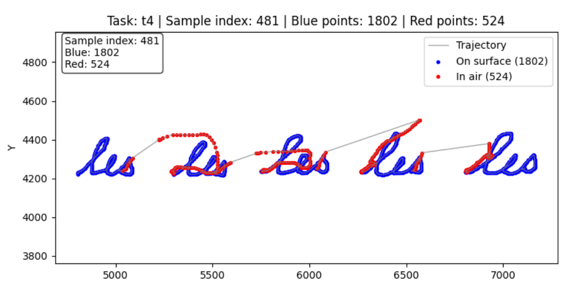


Figure 17: PaHaW task example: Task 4, repetitive cursive pattern.

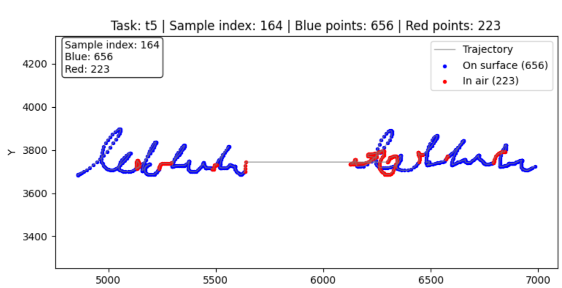


Figure 18: PaHaW task example: Task 5, word writing.

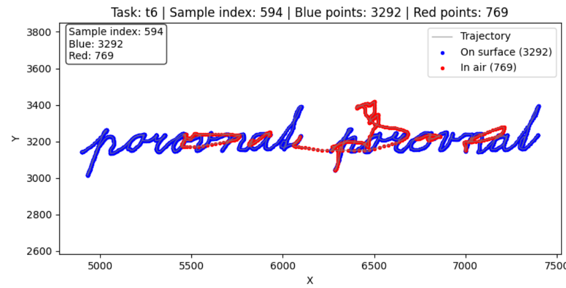


Figure 19: PaHaW task example: Task 6, word writing.

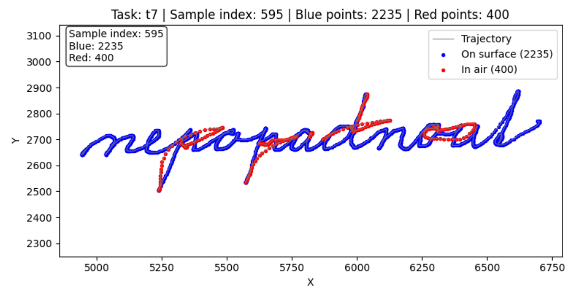


Figure 20: PaHaW task example: Task 7, word writing.

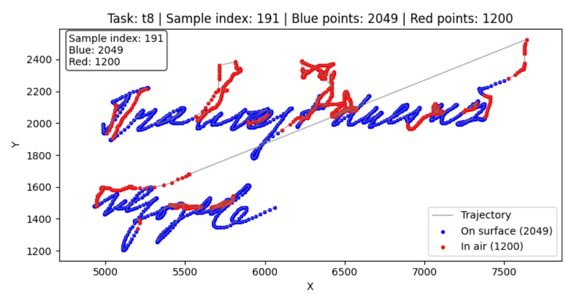


Figure 21: PaHaW task example: Task 8, word writing.

3.1.3. Fatigue

The Fatigue dataset is an online handwriting corpus acquired for fatigue assessment through pressure-related markers, including saturated pressure analysis [33]. From the perspective of the present study, it is relevant because it combines drawings, handwriting tasks, and signatures with markedly different temporal and pressure-related characteristics. This diversity makes it useful for examining watermarking behavior across signals with different lengths, task structures, and in-air versus on-surface compositions.

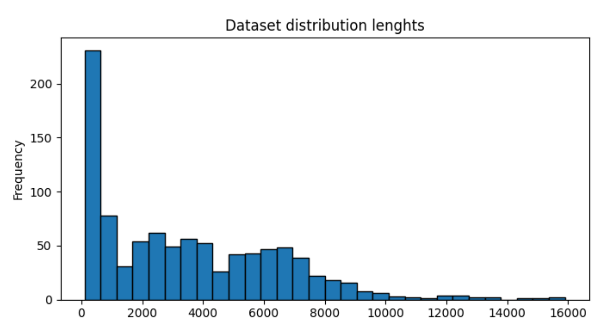


Figure 22: Fatigue dataset length distribution (frequency of samples vs. number of points).

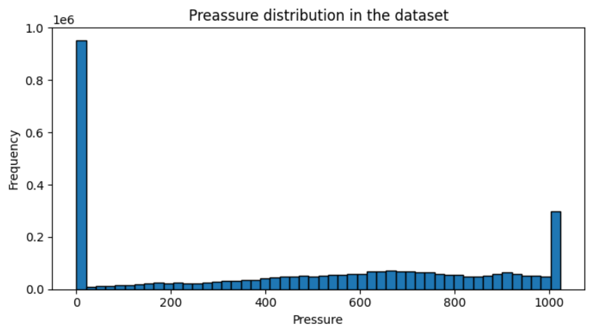
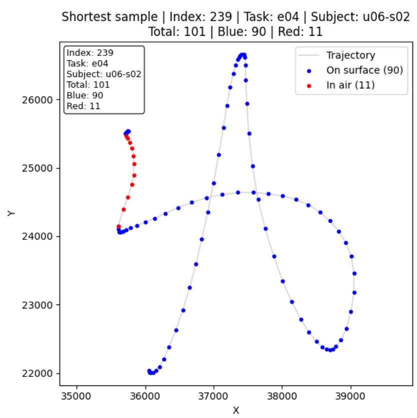
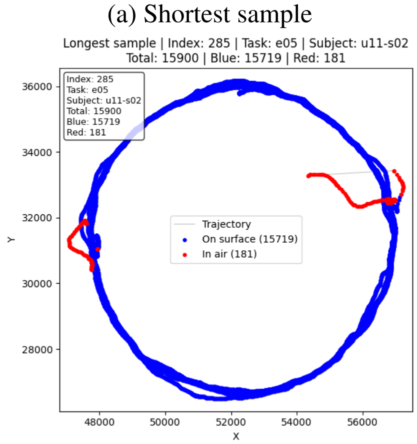


Figure 23: Fatigue dataset pressure distribution (frequency vs. pressure range).



(a) Shortest sample



(b) Longest sample

Figure 24: Fatigue dataset extreme-length samples: (a) shortest and (b) longest trajectories (in number of points).

A task-wise summary of trajectory length is provided in Table 5. The number of samples is highly consistent across tasks, with 105 samples per task in most cases and 106 for e08, while the average trajectory length varies substantially depending on the elicited content. The longest tasks in terms of mean length are e05 and e06, whereas the shortest are the signature tasks e04 and e08, followed by e07. These differences indicate that the dataset spans a broad range of temporal and structural demands, from short signatures to extended drawing and handwriting productions.

An additional aspect that makes the Fatigue dataset particularly relevant is its acquisition protocol. The dataset was collected from 21 participants under five different fatigue stages. Each participant performed the full set of tasks five times, once in each fatigue condition. This design was originally intended to study the relationship between online handwriting and the fatigue state of the writer, with particular interest in how fatigue may affect the temporal, spatial, and pressure-related properties of the recorded signal. From the perspective of the present work, this protocol also increases the structural diversity of the dataset, since the same tasks are available under repeated acquisitions associated with different writer conditions.

Table 5: Fatigue dataset summary by task: number of samples and length statistics (in points).

Task	Samples	Mean length	Min	Max
e01	105	2916.36	1000	8421
e02	105	5618.97	2542	13092
e03	105	2390.51	1303	5816
e04	105	497.68	101	1568
e05	105	7633.56	2188	15900
e06	105	6934.20	4607	10424
e07	105	520.89	258	1237
e08	106	484.59	102	1201
e09	105	4489.23	2788	9881

A task-wise summary of in-air and on-surface composition is provided in Table 6. Marked differences are observed across tasks. Some are almost entirely on-surface, such as e03, e05, and e07, all of which exceed 98.67% on-surface points. In contrast, tasks such as e02 and e06 contain a large proportion of in-air movements, with in-air percentages of 44.31% and 45.47%, respectively. Signature tasks e04 and e08 are much shorter and show similar distributions, with approximately 28% of points recorded in-air. Overall, these results indicate that pen-lift behavior and the balance between visible and non-visible trajectory segments depend strongly on task type, which is directly relevant to the watermarking analysis developed in this work.

Table 6: Fatigue dataset per-task mean number of points in-air ($p = 0$) and on-surface ($p > 0$), including their percentage over total points.

Task	Samples	Mean in-air	Mean on-surface	Mean total	% in-air	% on-surface
e01	105	1096.43	1819.93	2916.36	37.60	62.40
e02	105	2489.64	3129.33	5618.97	44.31	55.69
e03	105	8.14	2382.37	2390.51	0.34	99.66
e04	105	141.87	355.81	497.68	28.51	71.49
e05	105	18.94	7614.62	7633.56	0.25	99.75
e06	105	3153.10	3781.10	6934.20	45.47	54.53
e07	105	6.90	513.98	520.89	1.33	98.67
e08	106	133.74	350.86	484.59	27.60	72.40
e09	105	1791.52	2697.70	4489.23	39.91	60.09

The composition of the dataset is illustrated in Figures 25–33, which provide one representative example for each task. Following the acquisition protocol described in [33], the dataset contains nine tasks covering drawings, handwriting, and signatures: (e01) Folstein's pentagon copying, (e02) house drawing copying, (e03) Archimedes spiral drawing, (e04) signature, (e05) repeated concentric circles, (e06) words in capital letters copying, (e07) spring drawing, (e08) signature, and (e09) cursive sentence copying. This task inventory spans highly constrained geometric trajectories, short signatures, and longer handwriting productions, thereby providing a broad range of host-signal configurations.

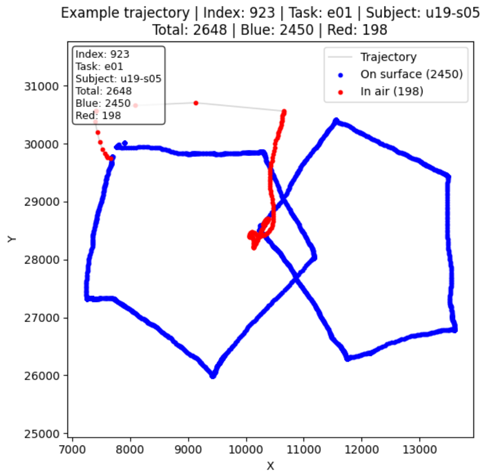


Figure 25: Fatigue dataset task example: (e01) Folstein's pentagon copying.

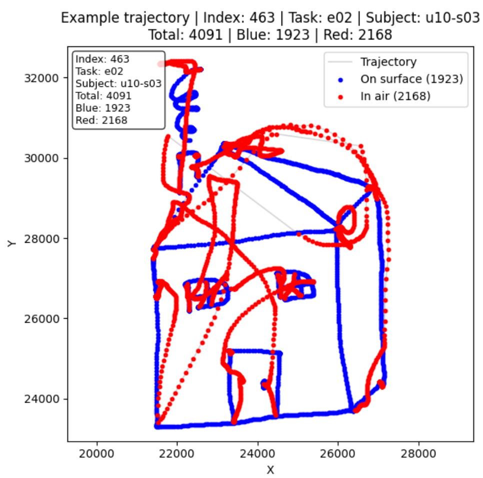


Figure 26: Fatigue dataset task example: (e02) house drawing copying.

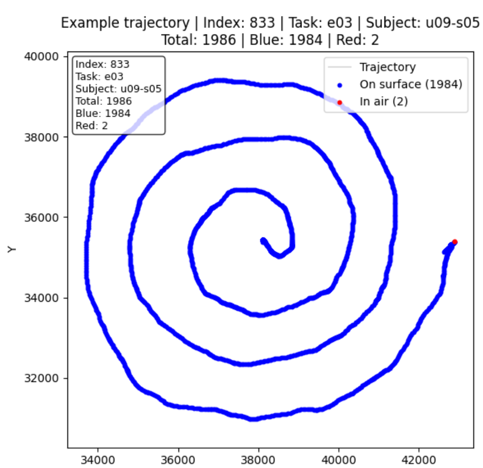


Figure 27: Fatigue dataset task example: (e03) Archimedes spiral drawing.

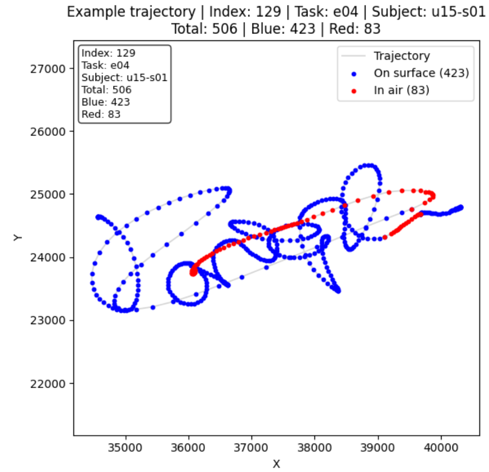


Figure 28: Fatigue dataset task example: (e04) signature.

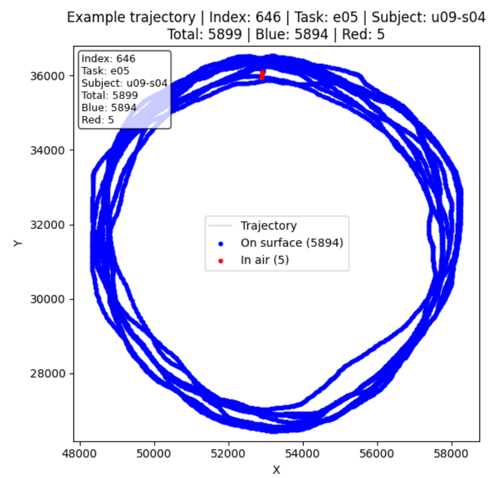


Figure 29: Fatigue dataset task example: (e05) repeated concentric circles.

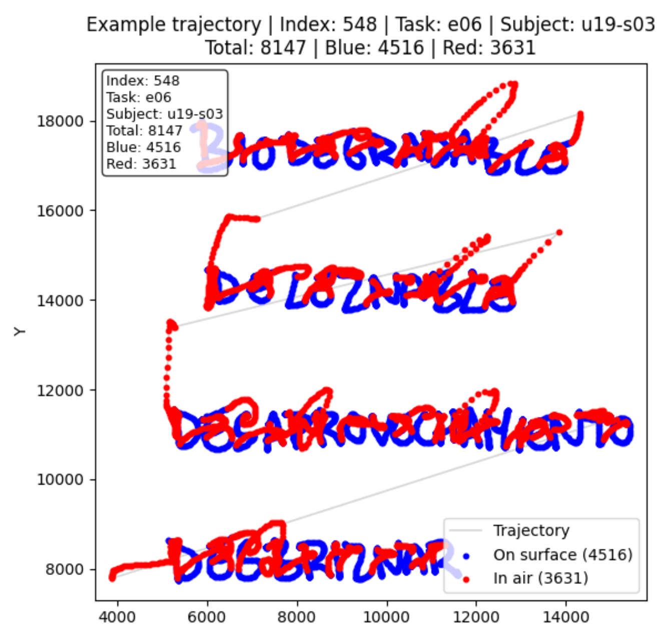


Figure 30: Fatigue dataset task example: (e06) words in capital letters copying.

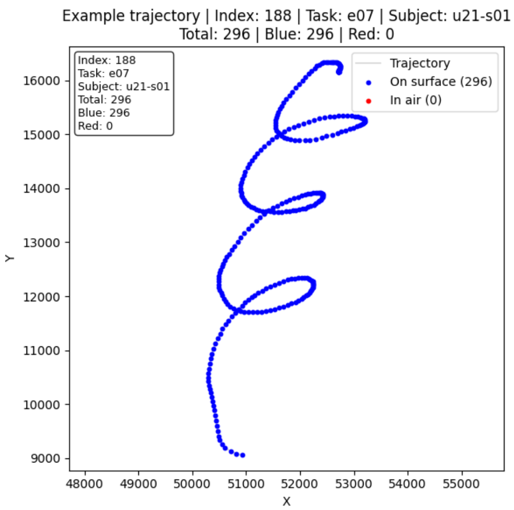


Figure 31: Fatigue dataset task example: (e07) spring drawing.

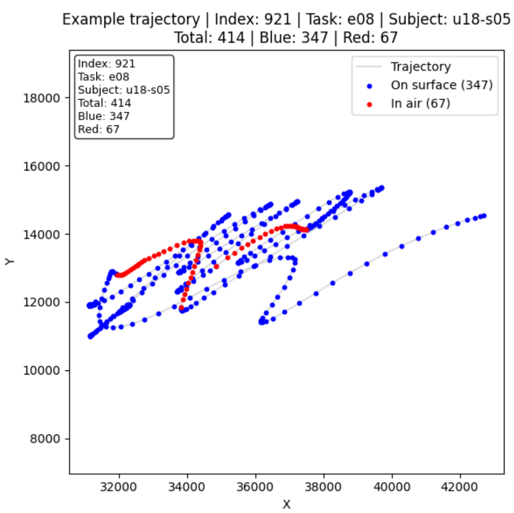


Figure 32: Fatigue dataset task example: (e08) signature.

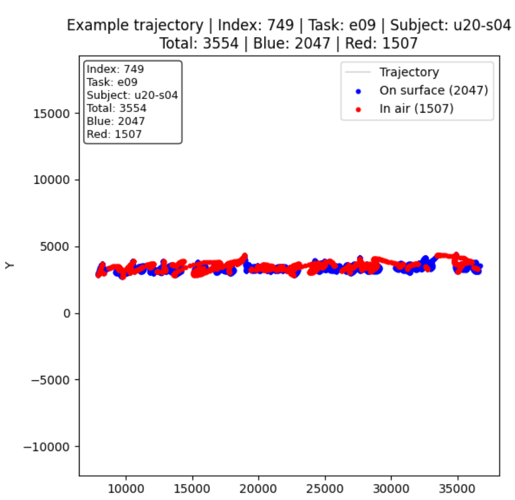


Figure 33: Fatigue dataset task example: (e09) cursive sentence copying.

3.1.4. PECT

PECT is an online handwriting dataset composed of drawing, pattern, handwriting, and signature tasks with markedly different structural properties [34]. From the perspective of the present study, it is especially relevant because it provides a broad range of trajectory lengths and strongly task-dependent in-air versus on-surface compositions, making it suitable for analyzing watermarking behavior across heterogeneous host signals.

An additional aspect that makes PECT particularly relevant is its original clinical motivation. The PECT dataset was acquired to study cognitive impairment in older adults through handwriting and drawing tasks performed on a digitizing tablet. In particular, the dataset was designed in a scenario where users were classified into healthy and cognitive-impairment groups based on the Pentagon Copying Test, making it especially relevant for the analysis of handwriting in an aging and clinically meaningful context.

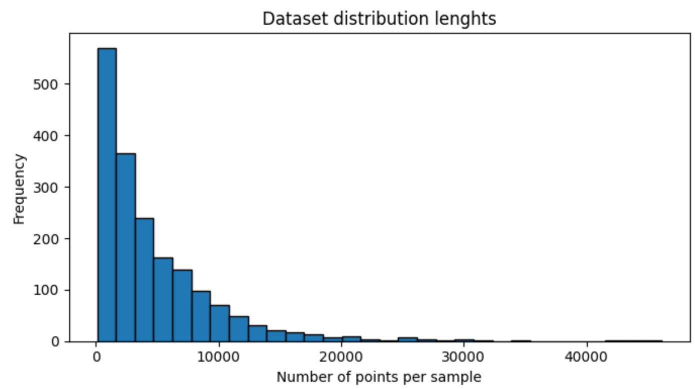


Figure 34: PECT length distribution (frequency of samples vs. number of points).

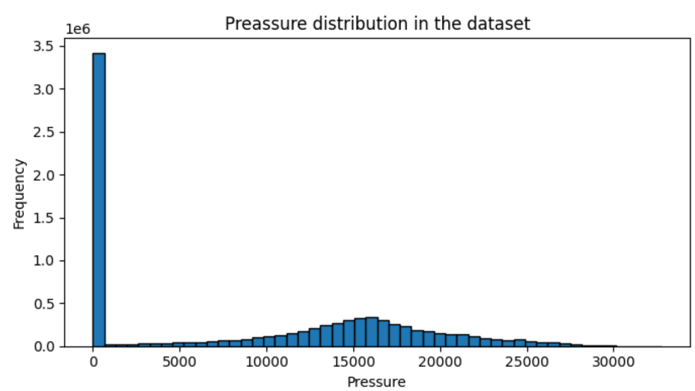
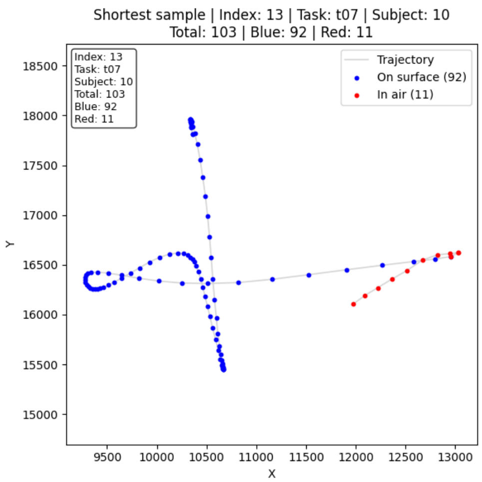
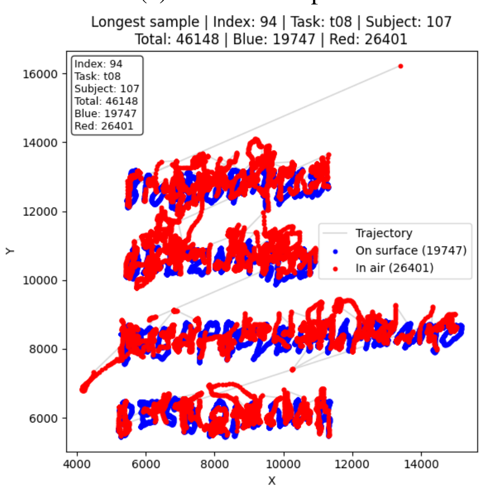


Figure 35: PECT pressure distribution (frequency vs. pressure range).

A task-wise summary of trajectory length is provided in Table 7. The dataset contains between 178 and 184 samples per task, and the mean trajectory length varies substantially across tasks. In particular, t08 and t09 are the longest tasks on average, with mean lengths of 13180.73 and 8607.94 points, respectively, whereas t07 and t10 are among the shortest, with mean lengths close to 1000 points. The maximum lengths also vary considerably, reaching 46148 points in t08 and 42986 in t09. These differences indicate that PECT covers tasks with markedly different temporal and structural demands, which is especially useful for comparative watermarking analysis.



(a) Shortest sample



(b) Longest sample

Figure 36: PECT extreme-length samples: (a) shortest and (b) longest trajectories (in number of points).

Table 7: PECT summary by task: number of samples and length statistics (in points).

Task	Samples	Mean length	Min	Max
t01	184	3549.64	1162	10004
t02	184	7837.26	3131	25560
t03	184	3424.71	1299	12110
t04	184	1664.91	480	3782
t05	184	1144.35	378	5841
t06	184	5927.38	1312	25193
t07	179	1049.22	103	3772
t08	179	13180.73	5210	46148
t09	178	8607.94	325	42986
t10	178	1111.31	170	4359

The task inventory is illustrated in Figures 37–46, which provide one representative example for each of the ten tasks included in the dataset. PECT combines geometric drawings, repetitive patterns, handwriting, and signature tasks: t01 corresponds to Folstein's pentagon copying, t02 of a house, t03 of an Archimedean spiral, t04 of a straight line, t05 corresponds to a spring drawing, t06 of circle loops, t07 of a signature, t08 of four words, t09 of a sentence, and t10 of a second signature task. This diversity is particularly relevant for the present study because it enables the analysis of watermarking behavior under a wide range of host-signal configurations, from highly constrained geometric traces to more natural handwritten and biometric productions.

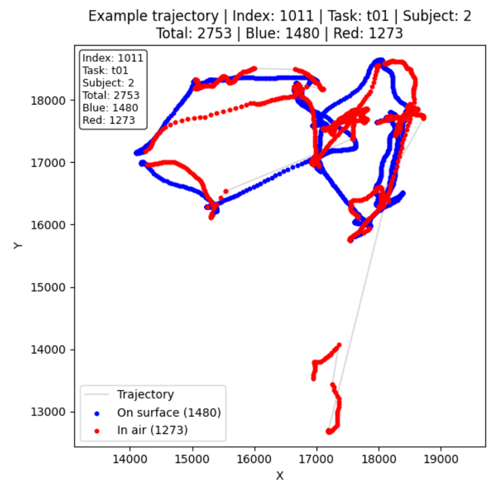


Figure 37: PECT task example: t01, Folstein's pentagon copying.

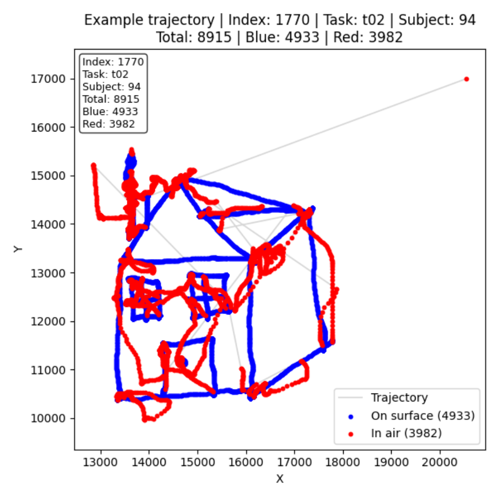


Figure 38: PECT task example: t02, a house.

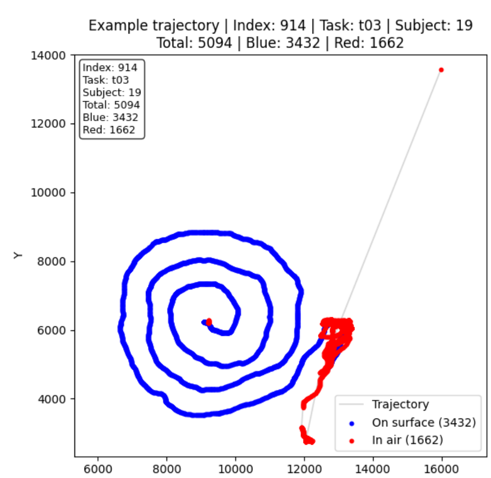


Figure 39: PECT task example: t03, Archimedean spiral.

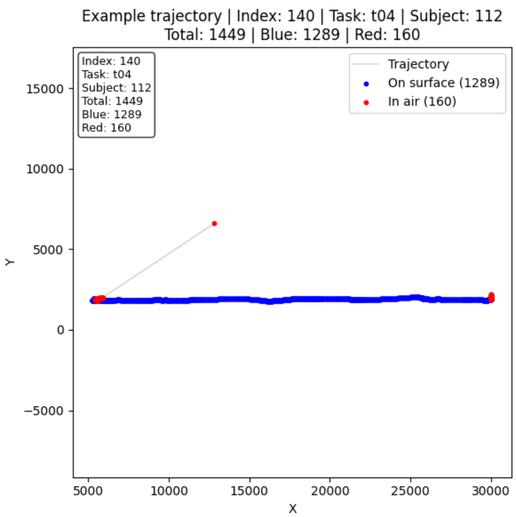


Figure 40: PECT task example: t04, straight line.

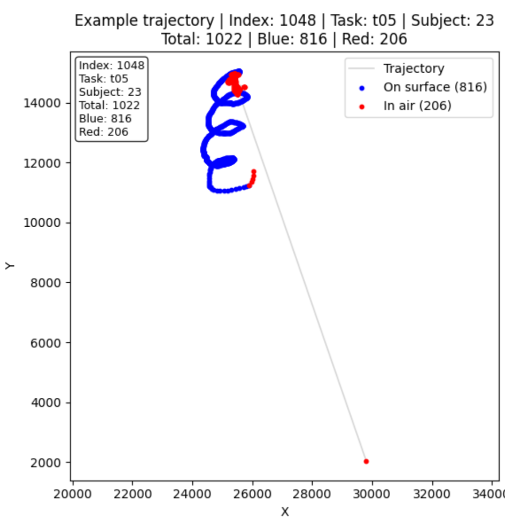


Figure 41: PECT task example: t05, spring drawing.

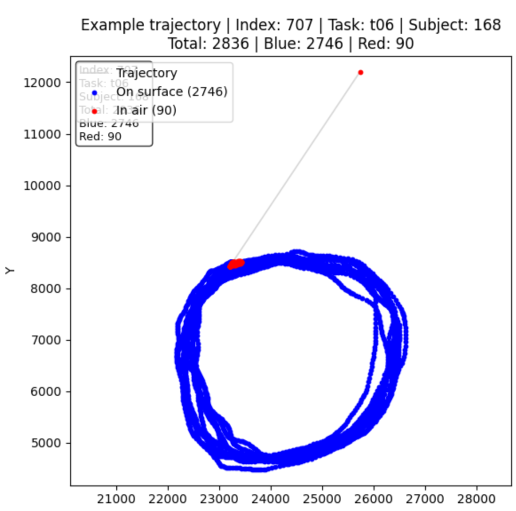


Figure 42: PECT task example: t06, circle loops.

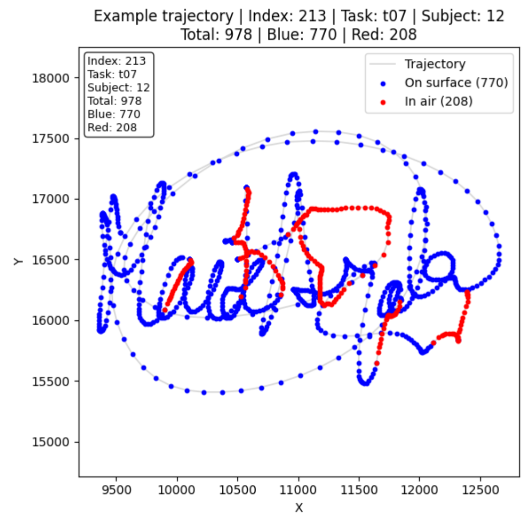


Figure 43: PECT task example: t07, signature.

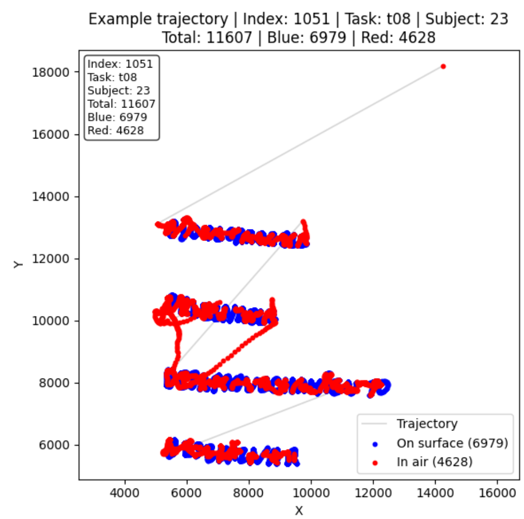


Figure 44: PECT task example: t08, four words.

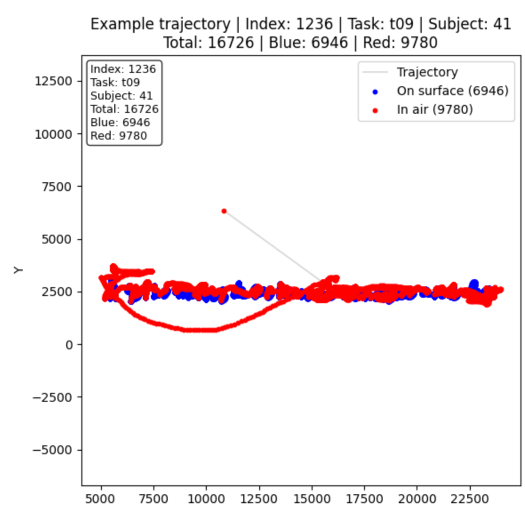


Figure 45: PECT task example: t09, sentence.

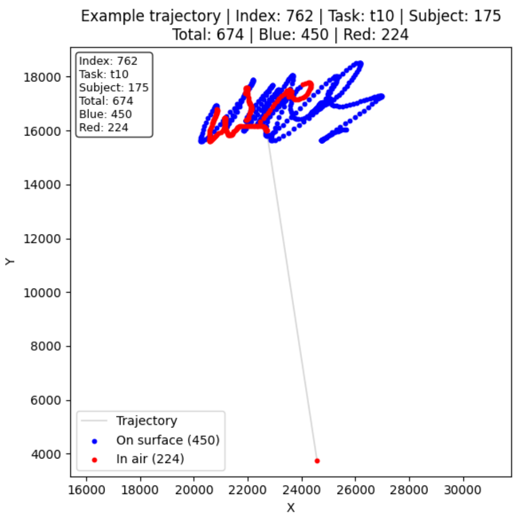


Figure 46: PECT task example: t10, signature.

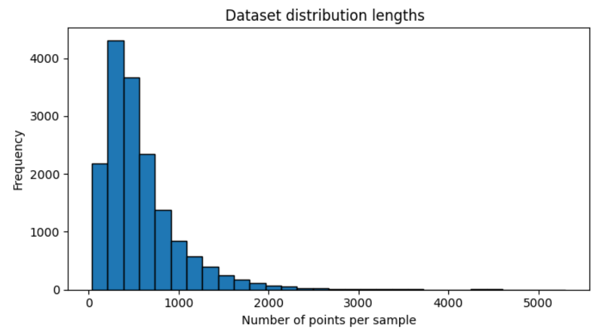


Figure 47: MCYT length distribution (frequency of signatures vs. number of points).

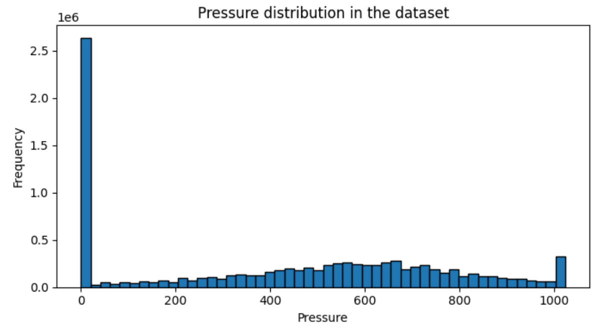


Figure 48: MCYT pressure distribution (frequency vs. pressure range).

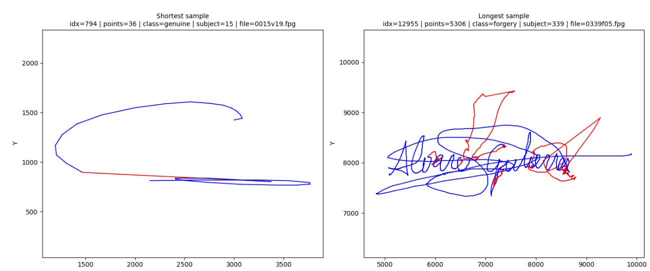


Figure 49: MCYT extreme-length signatures: shortest and longest samples (in number of points).

A task-wise summary of in-air and on-surface composition is provided in Table 8. The results show marked differences across tasks. Tasks t08 and t09, corresponding to four-word and sentence handwriting, exhibit the highest in-air proportions, reaching 52.01% and 50.87%, respectively, which indicates frequent pen lifts and long transition segments. In contrast, t06, corresponding to circle loops, is predominantly on-surface, with 91.69% of its points recorded in contact with the surface. The remaining tasks occupy intermediate positions, with in-air percentages ranging from 8.31% to 38.78%. Overall, these results confirm that PECT contains strongly task-dependent signal structures, which is directly relevant to the comparative watermarking analysis carried out in this work.

Table 8: PECT per-task mean number of points in-air ($p = 0$) and on-surface ($p > 0$), including their percentage over total points.

Task	Samples	Mean in-air	Mean on-surface	Mean total	% in-air	% on-surface
t01	184	1376.57	2173.08	3549.64	38.78	61.22
t02	184	3826.88	4010.38	7837.26	48.83	51.17
t03	184	608.79	2815.92	3424.71	17.78	82.22
t04	184	314.36	1350.55	1664.91	18.88	81.12
t05	184	287.05	857.29	1144.35	25.08	74.92
t06	184	492.84	5434.54	5927.38	8.31	91.69
t07	179	266.65	782.58	1049.22	25.41	74.59
t08	179	6855.43	6325.30	13180.73	52.01	47.99
t09	178	4378.76	4229.18	8607.94	50.87	49.13
t10	178	357.94	753.37	1111.31	32.21	67.79

3.1.5. MCYT

The on-line signature subset of MCYT is one of the most established benchmark resources for dynamic signature analysis under genuine and forgery conditions [35]. In the context of the present study, it is especially relevant because it provides a security-oriented signature scenario in which signal properties can be examined not only in terms of length and pressure usage, but also with respect to differences between authentic and forged samples. This makes MCYT particularly useful for studying watermarking behavior in host signals associated with realistic biometric variability.

A class-wise summary of in-air and on-surface composition is provided in Table 9. The comparison between genuine signatures and forgeries reveals a clear difference in trajectory composition: forgeries contain a higher proportion of in-air points (31.03%) than genuine signatures (20.80%). In absolute terms, forged samples also contain more in-air and on-surface points on average, which is consistent with their greater mean total length. This pattern suggests that forged signing attempts may involve more pen lifts and between-stroke movements than genuine signatures, a difference that is relevant not only for dynamic signature analysis but also for understanding how host-signal structure may vary across biometric classes.

Table 9: MCYT mean number of points in-air ($p = 0$) and on-surface ($p > 0$) for genuine signatures and forgeries, including percentages over total points.

Class	Samples	Mean in-air	Mean on-surface	Mean total	% in-air	% on-surface
Forgery	8250	222.80	495.13	717.92	31.03	68.97
Genuine	8250	92.52	352.40	444.92	20.80	79.20

3.1.6. BiosecurID

The on-line signature subset of BiosecurID is another widely used benchmark resource for dynamic signature analysis under genuine and forgery conditions [36]. In the context of the present study, it is relevant because it provides a complementary security-oriented signature scenario in which host-signal properties can be examined across authentic and forged samples. This makes BiosecurID particularly useful for analyzing watermarking behavior under realistic biometric variability.

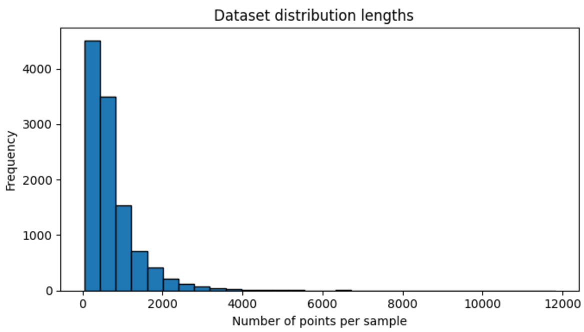


Figure 50: BiosecurID length distribution (frequency of signatures vs. number of points).

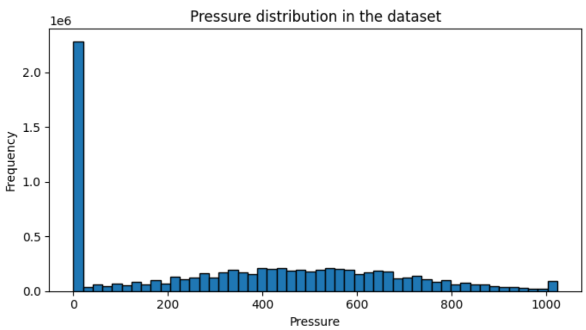


Figure 51: BiosecurID pressure distribution (frequency vs. pressure range).

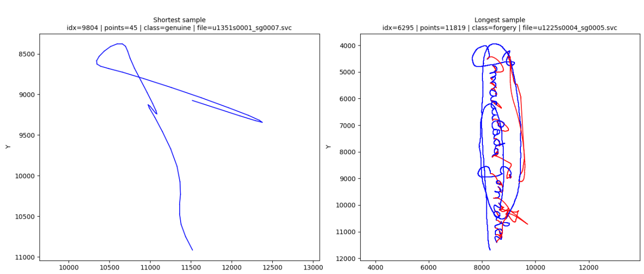


Figure 52: BiosecurID extreme-length signatures: shortest and longest samples (in number of points).

A class-wise summary of in-air and on-surface composition is provided in Table 10. As in MCYT, a clear difference is observed between genuine signatures and forgeries. Forged samples contain a higher proportion of in-air points (31.15%) than genuine signatures (21.96%), and they also exhibit larger mean numbers of both in-air and on-surface points, which is consistent with their greater mean total length. This pattern suggests that forged signing attempts may involve more pen lifts and longer between-stroke movements than genuine signatures, providing further evidence that host-signal structure can vary systematically across biometric classes.

Table 10: BiosecurID mean number of points in-air ($p = 0$) and on-surface ($p > 0$) for genuine signatures and forgeries, including percentages over total points.

Class	Samples	Mean in-air	Mean on-surface	Mean total	% in-air	% on-surface
Forgery	4798	328.43	725.82	1054.25	31.15	68.85
Genuine	6400	104.78	372.43	477.20	21.96	78.04

3.2. Task Taxonomy

Beyond the dataset-specific descriptions presented above, the analyzed material can be organized according to the type of task performed, since task structure directly influences the temporal and geometric properties of the resulting signal. This categorization is relevant in the present study because watermarking capacity is expected to depend not only on the embedding method, but also on the nature of the host trajectory.

The tasks considered in this work can be grouped into five broad categories: signatures, isolated words or short lexical items, extended text, geometric drawings, and repetitive patterns. Signature tasks are represented by the on-line signature subsets of MCYT and BiosecurID, as well as by Fatigue e04 and e08 and PECT t07 and t10. These signals are typically shorter than extended handwriting samples and are especially relevant because they arise in biometric settings involving genuine and, in some cases, forged productions.

A second category corresponds to isolated words or short handwriting productions. This group includes, for example, EMOTHAW t3, PaHaW t5–t8, Fatigue e06, and PECT t08. These tasks generally produce medium-length trajectories and show variable in-air activity depending on the amount of pen lifting and between-word transitions. Extended-text tasks form a separate category because they usually generate longer trajectories and richer temporal structure; representative examples include EMOTHAW t7, Fatigue e09, and PECT t09.

A third group comprises geometric drawings, such as houses, clocks, spirals, straight lines, springs, or pentagon copying tasks. Representative examples are EMOTHAW t1, t2, and t6; PaHaW t1; Fatigue e01, e02, e03, and e07; and PECT t01, t02, t03, t04, and t05. These tasks often generate structured trajectories with task-dependent variations in continuity and pressure usage. Finally, repetitive patterns define a distinct category that includes EMOTHAW t4 and t5, PaHaW t2–t4, Fatigue e05, and PECT t06. These signals are relevant because they combine regular geometric organization with task-dependent continuity and contact patterns.

This taxonomy provides a common interpretative framework for the host signals analyzed in this work. Rather than considering all recordings as equivalent handwriting traces, the proposed grouping makes it possible to relate watermarking behavior to structural differences in the underlying task.

3.3. Signal Descriptors Relevant to Watermarking

The dataset analysis presented in the previous subsections yields several signal descriptors that are directly relevant to watermark embedding. The first and most immediate descriptor is trajectory length, measured as the number of recorded points. This variable is important because it conditions the amount of available host signal and therefore the potential payload that can be embedded.

A second relevant descriptor is the composition of the trajectory in terms of in-air and on-surface points. In-air points correspond to samples recorded with $p = 0$, whereas on-surface points correspond to $p > 0$. Their absolute counts and relative percentages are important because they characterize how much of the trajectory corresponds to visible writing versus pen-lift and transition movements. From a watermarking perspective, this distinction is especially relevant because embedding on visible and non-visible segments may lead to different distortion and robustness behaviors.

A third descriptor is task category, as defined in the previous subsection. Signatures, isolated words, extended text, geometric drawings, and repetitive patterns differ not only in semantic content but also in temporal continuity, structural regularity, and typical point distribution. These differences are expected to affect watermarking performance even when the embedding method remains unchanged.

Finally, for MCYT and BiosecurID, class condition constitutes an additional descriptor. In these datasets, genuine signatures and forgeries exhibit different average trajectory lengths and different in-air versus on-surface proportions. This indicates that host-signal structure may vary not only across datasets and task categories, but also across biometric classes within the same acquisition setting.

In summary, the descriptors considered most relevant for the present work are mean trajectory length, minimum and maximum length, mean in-air points, mean on-surface points, in-air percentage, on-surface percentage, task category, and class condition when applicable. Together, these variables provide a compact characterization of the host signals and are used to interpret the comparative watermarking results reported in the following sections.

3.4. Summary of Datasets and Tasks

Tables 11–16 summarize the main characteristics of the datasets and task or class conditions analyzed in this work. The selected descriptors condense the most relevant properties of the host signals from the perspective of watermarking: sample availability, average trajectory length, relative amount of in-air and on-surface movement, and task category. This summary provides a compact view of the variability covered by the study and serves as a reference for the comparative analyses presented later.

Table 11: Summary of the analyzed tasks/classes for EMOTHAW. Length statistics are given in number of points.

Task/Class	Category	Samples	Mean length	% in-air
t1_data	Geometric drawing	129	2568.58	43.74
t2_data	Geometric drawing	129	4416.14	45.45
t3_data	Short text / words	129	3918.14	48.21
t4_data	Repetitive pattern	128	1378.93	2.13
t5_data	Repetitive pattern	129	1202.18	0.82
t6_data	Geometric drawing	129	3811.75	53.71
t7_data	Extended text	129	3290.97	38.34

Table 12: Summary of the analyzed tasks/classes for PaHaW. Length statistics are given in number of points.

Task/Class	Category	Samples	Mean length	% in-air
t1	Geometric drawing	72	2873.69	2.21
t2	Repetitive pattern	75	1668.01	32.89
t3	Repetitive pattern	75	1984.16	24.59
t4	Repetitive pattern	75	2305.01	22.76
t5	Short text / words	75	2608.48	18.37
t6	Short text / words	75	2315.08	21.44
t7	Short text / words	75	1469.29	17.28
t8	Short text / words	75	3086.13	34.91

Table 13: Summary of the analyzed tasks/classes for Fatigue. Length statistics are given in number of points.

Task/Class	Category	Samples	Mean length	% in-air
e01	Geometric drawing	105	2916.36	37.60
e02	Geometric drawing	105	5618.97	44.31
e03	Geometric drawing	105	2390.51	0.34
e04	Signature	105	497.68	28.51
e05	Repetitive pattern	105	7633.56	0.25
e06	Short text / words	105	6934.20	45.47
e07	Geometric drawing	105	520.89	1.33
e08	Signature	106	484.59	27.60
e09	Extended text	105	4489.23	39.91

Table 14: Summary of the analyzed tasks/classes for PECT. Length statistics are given in number of points.

Task/Class	Category	Samples	Mean length	% in-air
t01	Geometric drawing	184	3549.64	38.78
t02	Geometric drawing	184	7837.26	48.83
t03	Geometric drawing	184	3424.71	17.78
t04	Geometric drawing	184	1664.91	18.88
t05	Repetitive-pattern	184	1144.35	25.08
t06	Repetitive pattern	184	5927.38	8.31
t07	Signature	179	1049.22	25.41
t08	Short text / words	179	13180.73	52.01
t09	Extended text	178	8607.94	50.87
t10	Signature	178	1111.31	32.21

Table 15: Summary of the analyzed tasks/classes for MCYT. Length statistics are given in number of points.

Task/Class	Category	Samples	Mean length	% in-air
Genuine	Signature	8250	444.92	20.80
Forgery	Signature	8250	717.92	31.03

Table 16: Summary of the analyzed tasks/classes for BiosecurID. Length statistics are given in number of points.

Task/Class	Category	Samples	Mean length	% in-air
Genuine	Signature	6400	477.20	21.96
Forgery	Signature	4798	1054.25	31.15

4. Watermarking Methods Under Comparison

4.1. Method Selection Criteria

The methods considered in this work were selected to provide a representative comparison across the main watermarking families applicable to online handwriting signals. Rather than focusing on variants of a single embedding strategy, the objective is to include methods that differ in embedding domain, reversibility, robustness, and expected payload behavior. This makes it possible to analyze how the main trade-offs between payload, distortion, and robustness vary not only across datasets and tasks, but also across fundamentally different watermarking paradigms.

Accordingly, the comparison includes a spatial-domain method (LSB), a reversible difference-based method (DE), a quantization-based method (QIM), and four transform-domain methods (DCT, DWT, DFT, and SVD). This selection is consistent with the broader watermarking literature, where spatial and transform techniques are commonly regarded as the two main design families, while reversible and quantization-based methods define relevant subgroups with distinct operational properties [21, 22, 23, 24].

The selected set also reflects the methods already explored in the context of online signature and related signal embedding problems. In particular, DCT-based watermarking and blind DCT extraction have been studied for online signature signals [10, 11], while LSB and Difference Expansion have been analyzed as spatial-domain strategies for the same modality [12]. The remaining methods were included because they represent widely used watermarking paradigms with well-known robustness or compactness properties in signal and multimedia applications [25, 28, 29, 15].

4.2. LSB-Based Watermarking

Least Significant Bit (LSB) watermarking is a spatial-domain embedding strategy in which the watermark is inserted directly into the least significant bits of the host signal samples. In the present study, this approach is applied to the coordinate sequences of the online handwriting signal. Its main advantage is simplicity, together with a potentially high embedding capacity when multiple least significant bits are modified. However, because the embedding takes place directly in the signal domain, LSB-based methods are generally expected to be sensitive to perturbations such as noise and filtering [12, 21, 22].

4.3. Difference Expansion Watermarking

Difference Expansion (DE), originally introduced by Tian, is a reversible data-hiding technique that embeds information by expanding the difference between pairs of host-signal samples [37]. Exact reversibility requires the selected pairs to satisfy the expandability and coordinate-range constraints of the DE transform [37]. In online handwriting, this strategy can be applied to consecutive coordinate values, allowing watermark insertion while preserving the possibility of exact recovery of the original signal in the absence of attacks. This reversibility makes DE attractive when signal preservation is important, although it is generally more fragile under common signal processing perturbations than robust transform-domain approaches [12, 19].

4.4. QIM-Based Watermarking

Quantization Index Modulation (QIM) embeds information by selecting a quantizer indexed by the watermark symbol and quantizing the host sample using the corresponding quantization lattice. Its main advantage is that the embedding strength can be controlled explicitly through the quantization step, which provides a direct way to regulate the robustness–distortion compromise [38]. In comparative studies, QIM is particularly useful because it represents a distinct design philosophy, intermediate between direct sample modification and transform-domain modulation [28]. For online handwriting signals, this makes it a relevant reference point for capacity and robustness analysis.

4.5. DCT-Based Watermarking

Discrete Cosine Transform (DCT)-based watermarking embeds information in a transformed representation of the host signal by modifying selected frequency-domain coefficients. In online signature signals, DCT-based formulations have already been explored in both single-bit and multi-bit settings, as well as in blind extraction scenarios [10, 11]. Because embedding is performed in the transform domain, DCT methods are generally expected to provide a more favorable robustness–transparency balance than purely spatial approaches, although the achieved payload depends on the embedding configuration and the structure of the host signal.

4.6. DWT-Based Watermarking

Discrete Wavelet Transform (DWT)-based watermarking embeds the watermark into a multiresolution representation of the host signal, typically by modifying selected wavelet detail subbands at one or more decomposition levels [39]. This strategy is attractive because it combines frequency localization with scale decomposition, which often leads to a good balance between imperceptibility and robustness [15, 30]. In the present comparison, DWT represents a transform-domain approach in which watermarking behavior may differ from DCT and DFT due to the hierarchical structure of the wavelet representation.

4.7. DFT-Based Watermarking

Discrete Fourier Transform (DFT)-based watermarking embeds information in the spectral representation of the host signal, typically through modifications of Fourier magnitude coefficients [40]. As in other transform-domain approaches, the rationale is that properly selected spectral components may offer increased robustness against common signal processing operations [25]. In the context of online handwriting, DFT provides an alternative frequency-domain strategy whose behavior can be compared with DCT and DWT in terms of payload, distortion, and robustness.

4.8. SVD-Based Watermarking

Singular Value Decomposition (SVD)-based watermarking embeds information through the singular-value matrix of a matrix representation of the host signal [41]. The main interest of this approach lies in the stability of singular values under a range of perturbations, which often yields good robustness properties [29, 25].

In the online-handwriting implementation considered here, SVD-based embedding provides fewer embedding locations than sample-wise or coefficient-wise methods, resulting in a lower payload but potentially favorable robustness.

4.9. Comparative Summary of the Methods

Table 17 summarizes the main methodological characteristics of the watermarking approaches considered in this work. Performance-related comparisons are presented in the following sections.

Table 17: Methodological characteristics of the evaluated watermarking methods.

Method	Embedding configuration	Rev.
LSB	Spatial domain; coordinate samples; b least significant bits	No
DE	Difference-based; coordinate pairs; payload rate ρ	Yes[a]
QIM	Quantization-based; coordinate samples; step Δ	No
DCT	Transform domain; all DCT coefficients; gain α	No
DWT	Transform domain; approximation coefficients; gain α	No
DFT	Transform domain; spectral magnitudes; gain α	No
SVD	Transform domain; singular values; gain α	No

[a] Exact host recovery is obtained in the absence of attacks when the reversibility conditions of the DE implementation are satisfied.

The configurations summarized in Table 17 describe the embedding domain, control parameter, and reversibility of the evaluated approaches. Extraction requirements are implementation-dependent and are therefore not treated as intrinsic properties of the corresponding watermarking families.

5. Experimental Framework

5.1. Signal Representation

The host signals considered in this work are online handwriting trajectories represented by their coordinate sequences and associated pressure values. In practical terms, each sample is modeled through its X and Y coordinate sequences together with the pressure signal, which allows the distinction between in-air points ($p = 0$) and on-surface points ($p > 0$). This distinction is relevant because the full trajectory includes both visible pen strokes and non-contact pen movements, and these two components may exhibit different watermarking behavior.

Although the controlled comparison in this work uses the coordinate sequences as the common embedding baseline, recent pressure-domain fragile watermarking work shows that the pressure channel can also act as an integrity carrier, provided that zero-pressure pen-up samples are strictly preserved [13].

Accordingly, the analysis considers the complete recorded trajectory as the host signal, while also taking into account the relative proportion of on-surface and in-air points as a structural descriptor. This representation is consistent across the datasets described in the previous section and provides a common basis for method comparison.

5.2. Reference Signature Selection and Characterization

The controlled distortion and robustness experiments were conducted using a genuine online signature voluntarily produced by the first author with a Wacom Intuos Pro pen tablet. The signature was acquired specifically for this study and is independent of the datasets characterized in Section 3. Therefore, no external participant or third-party biometric sample was involved in the controlled experimental comparison. The sample was selected to provide a fixed host signal under fully controlled conditions. Its purpose was not to represent the complete variability of online handwriting, but to isolate method-dependent differences by applying all watermarking approaches to the same coordinate trajectory and under the same attack and evaluation conditions. Table 18 summarizes the main structural characteristics of the reference signature. The sample contains 4004 recorded points, corresponding to 8008 coordinate values when the X and Y sequences are considered jointly. Of these samples, 1402 (35.01%) correspond to in-air movement and 2602 (64.99%) correspond to on-surface writing. The trajectory contains seven continuous pen-down strokes and six in-air segments.

Table 18: Structural characteristics of the reference online signature used in the controlled experiments. Coordinate measurements are expressed in acquisition units.

Characteristic	Value
Acquisition device	Wacom Intuos Pro
Total recorded points	4004
Total X and Y coordinate values	8008
In-air points ($p = 0$)	1402 (35.01%)
On-surface points ($p > 0$)	2602 (64.99%)
Continuous pen-down strokes	7
In-air segments	6
X-coordinate range	60888
Y-coordinate range	13879
Bounding-box diagonal	62449.78

Before watermark embedding, the X and Y coordinate sequences were translated so that the trajectory centroid coincided with the fixed reference point $(6350, 9700)$. This translation changes the absolute position of the signature but preserves its geometry, trajectory length, and in-air/on-surface composition. Using a single controlled sample facilitates a direct method-level comparison by holding writer identity, task type, acquisition device, and trajectory structure constant across all watermarking methods. Nevertheless, the results obtained from this signature cannot be generalized to other writers, datasets, devices, or handwriting tasks without additional experimental validation. Accordingly, the multidataset component of this study provides structural host-signal characterization, whereas the distortion and robustness experiments provide controlled reference-level evidence.

5.3. Embedding Configuration and Parameter Rationale

For each watermarking method, three configurations were evaluated to represent low, medium, and high embedding conditions. Since the compared techniques belong to different watermarking families, their control parameters are not directly equivalent. The selected values were therefore intended to provide three clearly differentiated operating points within each method, rather than to enforce identical payload or identical distortion across methods.

For DCT-, DWT-, DFT-, and SVD-based watermarking, the embedding gain was set to $\alpha \in \{3, 15, 75\}$. The same gain sequence was used across these transform-domain methods to maintain a consistent experimental structure and to examine the effect of progressively stronger coefficient modification. However, because the transformed representations have different numerical scales and embedding supports, the same value of α does not imply an equivalent distortion level across methods. For QIM, the quantization step was set to $\Delta \in \{2, 6, 14\}$. These values define progressively wider quantization intervals, providing a transition from low-distortion embedding to stronger quantization and potentially more stable watermark extraction. For LSB, the number of modified least significant bits was set to $b \in \{1, 3, 5\}$. This configuration provides a direct progression in both nominal payload and sample modification while keeping the embedding restricted to the lower-order bits of the coordinate values. For Difference Expansion, payload rates of $\rho \in \{0.25, 0.50, 1.00\}$ were considered. These values correspond to embedding in approximately one quarter, one half, and all of the available coordinate pairs, respectively. This makes it possible to examine how increasing the proportion of marked pairs affects payload, reversibility, distortion, and robustness. Table 19 summarizes the evaluated configurations and their experimental rationale.

Table 19: Embedding parameters and rationale for the low-, medium-, and high-strength configurations.

Method	Evaluated values	Rationale
DCT, DWT, DFT, SVD	$\alpha = \{3, 15, 75\}$	Common gain sequence providing progressively stronger transform-domain modification.
QIM	$\Delta = \{2, 6, 14\}$	Progressively wider quantization intervals to examine the distortion–robustness trade-off.
LSB	$b = \{1, 3, 5\}$ bits	Progressive increase in nominal payload and lower-order coordinate modification.
DE	$\rho = \{0.25, 0.50, 1.00\}$	Embedding in one quarter, one half, and all available coordinate pairs.

The parameter sets were not individually optimized to maximize the performance of each method. Instead, they were selected to support a controlled within-method analysis of increasing embedding intensity. Consequently, the low-, medium-, and high-strength labels indicate relative levels within each watermarking technique and should not be interpreted as equal-payload, equal-distortion, or otherwise numerically equivalent operating points across different methods.

5.4. Implementation and Reproducibility Details

All watermarking experiments were implemented in MATLAB R2024b using the Signal Processing Toolbox and Wavelet Toolbox functions required by the corresponding methods. Each technique was applied to the X and Y coordinate sequences of the same preprocessed reference signature. The only exception was SVD, which was applied directly to the $N \times 2$ matrix formed by the two coordinate sequences.

To ensure reproducibility, the pseudo-random number generator was initialized with `rng(0)` at the beginning of each method. Binary watermark sequences were generated using MATLAB pseudo-random functions, whereas the SVD watermark was represented by two bipolar values in $\{-1, +1\}$. For DCT, QIM, DWT, DFT, and SVD, the same generated watermark was used across the three embedding-strength configurations of each method. In LSB and DE, the watermark depended on the corresponding bit-depth or payload-rate configuration. Table 20 summarizes the embedding operation and the recovery information used by the specific implementations evaluated in the controlled experiments. These requirements are implementation-dependent and should not be interpreted as intrinsic properties of the corresponding watermarking families.

Table 20: Implementation details of the specific watermarking configurations evaluated in the controlled experiments.

Method	Implementation	Recovery data
LSB	Each coordinate was mapped to a multiple of 2^b using the MATLAB `fix` function, and a watermark symbol in the range 0 to $2^b - 1$ was then added. For non-negative coordinates, this operation is equivalent to replacing the b least significant bits. The watermark symbol was recovered using the remainder modulo 2^b.	Watermarked coordinates and bit depth b.
DE	Consecutive coordinate samples were grouped into pairs. A pseudo-random subset of pairs was selected according to the payload rate, and one bit was inserted by expanding each selected difference.	Side information identifying the marked pairs.
QIM	One binary bit was inserted into each coordinate value using two scalar quantization lattices separated by $\Delta/2$.	Watermarked coordinates and quantization step Δ.
DCT	A binary watermark was added to the DCT coefficients of the X and Y sequences using the embedding gain α.	Original DCT coefficients and α.
DWT	A one-level Haar wavelet decomposition was applied separately to X and Y, and the watermark was added to the approximation coefficients using α.	Original coordinate sequence, α, wavelet type, and padding information.
DFT	The watermark was added to the magnitudes of the positive-frequency coefficients, excluding the DC and Nyquist components. The original phase was preserved and conjugate symmetry was enforced.	Original spectral magnitudes and α.
SVD	The signature was represented as an $N \times 2$ matrix, and two bipolar watermark values were added to its two singular values using α.	Original singular values and α.

The low-pass attack was implemented in MATLAB as a fourth-order FIR filter designed with `fir1` and applied independently to the X and Y coordinate sequences using zero-phase filtering with `filtfilt`. Watermark estimates were converted to their corresponding binary representation before computing BER and NC. These metrics were calculated separately for the X and Y coordinate sequences for all methods except SVD. For SVD, BER and NC were computed over the two embedded singular-value watermark symbols. In the experimental summary, the coordinate-wise results were averaged for the non-SVD methods, whereas the joint two-symbol SVD result was reported directly.

5.5. Payload Definition

Payload is treated in this work as the amount of information embedded into the host signal under a given watermarking configuration. Because the analyzed datasets differ substantially in trajectory length and structure, payload must be interpreted both in absolute and relative terms.

The main payload quantity reported in this study is the nominal number of embedded bits under each stated configuration. Dataset- and task-level sample counts are analyzed separately as indicators of the available host-signal support. In particular, relative payload can be interpreted with respect to the total number of recorded points and, when relevant, with respect to the number of on-surface points. This distinction is useful because some signals contain large in-air segments, whereas others are almost entirely on-surface, and these differences may affect the practical embedding capacity of the host trajectory.

5.6. Evaluation Metrics

To quantify geometric distortion and watermark recovery after signal-processing attacks, three metrics were reported: normalized root mean square error (NRMSE), bit error rate (BER), and normalized correlation (NC). NRMSE quantifies the geometric distortion introduced in the online signature trajectory. BER and NC provide two complementary representations of watermark recovery: BER expresses the proportion of incorrectly recovered bits, whereas NC expresses the agreement between the original and recovered watermark in bipolar correlation form. Their joint reporting also provides an internal consistency check on the recovery results.

5.6.1. Normalized Root Mean Square Error (NRMSE)

The distortion between the original online signature and its watermarked version was measured by the normalized root mean square error computed on the two-dimensional trajectory. Let

$$\mathbf{s}_i = (x_i, y_i), \qquad i = 1, 2, \ldots, N$$

be the samples of the original signature, and let

$$\hat{\mathbf{s}}_i = (\hat{x}_i, \hat{y}_i)$$

be the corresponding samples of the watermarked signature. The pointwise squared error is defined as

$$d_i^2 = (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2.$$

The root mean square error is then given by

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2\right]}.$$

To make this measure independent of the absolute spatial scale of the signature, the RMSE was normalized by the diagonal of the bounding box of the original trajectory. If

$$\Delta x = \max_i x_i - \min_i x_i, \qquad \Delta y = \max_i y_i - \min_i y_i,$$

then the bounding-box diagonal is

$$D = \sqrt{(\Delta x)^2 + (\Delta y)^2}.$$

Finally, the normalized root mean square error is computed as

$$\mathrm{NRMSE} = \frac{\mathrm{RMSE}}{D}.$$

A lower NRMSE indicates lower geometric distortion and therefore better preservation of the original signature trajectory after watermark embedding.

5.6.2. Bit Error Rate (BER)

The bit error rate was used to quantify the proportion of watermark bits that were incorrectly recovered after applying a robustness attack, such as additive Gaussian noise or low-pass filtering. Let

$$\mathbf{w} = [w_1, w_2, \ldots, w_M]$$

denote the original embedded watermark bitstream, and let

$$\hat{\mathbf{w}} = [\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_M]$$

denote the extracted bitstream after the attack, where each element belongs to $\{0, 1\}$.

The BER is defined as

$$\mathrm{BER} = \frac{1}{M}\sum_{j=1}^{M} \mathbf{1}(w_j \neq \hat{w}_j),$$

where $\mathbf{1}(\cdot)$ is the indicator function, equal to 1 when the condition is true and 0 otherwise.

Equivalently, BER represents the fraction of erroneous bits with respect to the total number of embedded bits. A BER value of 0 indicates perfect recovery, whereas larger values indicate increasing degradation of the watermark under attack.

5.6.3. Normalized Correlation (NC)

The similarity between the original and extracted watermark was also assessed using normalized correlation. For binary watermark sequences, the bit values were first mapped into bipolar form:

$$0 \rightarrow -1, \qquad 1 \rightarrow +1.$$

Thus, given the original binary watermark $\mathbf{w}$ and the extracted watermark $\hat{\mathbf{w}}$, their bipolar representations are

$$\mathbf{b} = [b_1, b_2, \ldots, b_M], \qquad \hat{\mathbf{b}} = [\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_M],$$

with $b_j, \hat{b}_j \in \{-1, +1\}$.

The normalized correlation is then computed as

$$\mathrm{NC} = \frac{\sum_{j=1}^{M} b_j \hat{b}_j}{\sqrt{\left(\sum_{j=1}^{M} b_j^2\right)\left(\sum_{j=1}^{M} \hat{b}_j^2\right)}}.$$

Since the bipolar values satisfy $b_j^2 = 1$, the denominator acts as a normalization factor, ensuring that the metric is bounded. In practice, NC values close to 1 indicate high similarity between the original and recovered watermark, whereas values close to 0 indicate weak similarity. Negative values indicate strong disagreement between both sequences.

5.6.4. Interpretation in This Study

In the reported results, lower NRMSE values indicate better geometric preservation of the host trajectory. For robustness, lower BER and higher NC values indicate more reliable watermark recovery after attack.

BER and NC are therefore interpreted jointly as error-rate and correlation-based representations of the same recovery outcome, with their agreement providing an internal consistency check.

5.7. Attacks and Perturbations

Two perturbations are considered in the present study: additive Gaussian noise and low-pass filtering. These attacks were selected because they represent common signal alterations affecting dynamic handwriting trajectories and provide a simple but informative test of watermark robustness. Gaussian noise evaluates the sensitivity of the embedded watermark to additive coordinate-domain perturbations, whereas low-pass filtering evaluates the effect of smoothing and attenuation of high-frequency components on watermark recovery.

Let (X_w, Y_w) denote the watermarked coordinate sequences. Under Gaussian perturbation, the attacked signal is defined as

$$X_{\mathrm{noise}} = X_w + \sigma_{\mathrm{noise}}\varepsilon_X, \qquad Y_{\mathrm{noise}} = Y_w + \sigma_{\mathrm{noise}}\varepsilon_Y, \tag{1}$$

where ε_X and ε_Y are independent standard Gaussian random variables and $\sigma_{\mathrm{noise}} = 4$.

Low-pass filtering is applied to the watermarked coordinate sequences as

$$X_{\mathrm{lp}} = \mathcal{F}_{\mathrm{LP}}(X_w), \qquad Y_{\mathrm{lp}} = \mathcal{F}_{\mathrm{LP}}(Y_w), \tag{2}$$

where $\mathcal{F}_{\mathrm{LP}}$ denotes a low-pass filtering operator with normalized cutoff frequency $f_c = 0.99$ and filter order 4.

The Gaussian perturbations were generated independently for the X and Y coordinate sequences using zero-mean, unit-variance Gaussian samples scaled by $\sigma_{\mathrm{noise}} = 4$. The low-pass attack was implemented as a fourth-order FIR filter designed with the MATLAB function `fir1` and applied using zero-phase forward–reverse filtering with `filtfilt`. In both cases, the resulting coordinate values were rounded to the nearest integer before watermark extraction. The same attack configuration was applied to all watermarking methods and embedding-strength settings.

5.8. Analysis Dimensions

The comparative analysis is carried out along several complementary dimensions. First, methods are compared in order to identify differences associated with the embedding family itself. Second, datasets and task categories are considered in order to examine how host-signal structure is associated with watermarking behavior. Third, when applicable, class-dependent differences are also taken into account, particularly in the case of genuine and forged signatures.

Overall, the experimental framework combines heterogeneous host-signal characterization with a controlled reference-level comparison of watermarking methods, with particular attention to the interaction between payload, distortion, and robustness.

To avoid conflating different forms of evidence, the results are described according to their origin. Dataset statistics and the reported NRMSE, BER, and NC values are measured directly from the analyzed records and controlled experiments. Nominal payload values or analytical expressions are derived from the stated implementation configurations. Statements concerning the expected influence of trajectory length, task type, or in-air/on-surface composition are presented as interpretations of the observed signal structure and not as experimentally validated causal relationships.

All methods were evaluated using the same reference signature, preprocessing procedure, attack parameters, and metric definitions. The same NRMSE definition was used to quantify geometric distortion, whereas BER and NC were used to assess watermark recovery whenever direct comparison was applicable. Reversibility was reported separately because it is a method-specific property rather than a common continuous metric. The low-, medium-, and high-strength configurations represent relative operating points within each method and should not be interpreted as equal-payload or equal-distortion conditions across different methods.

6. Descriptive Analysis of the Signals

6.1. Length Variability Across Datasets and Tasks

The datasets analyzed in this work exhibit substantial variability in trajectory length, both across datasets and within the task structure of each corpus. As shown in the dataset descriptions and in Tables 11–16, the mean number of recorded points ranges from short signature-like signals with fewer than 500 points, such as Fatigue e04 and e08, to extended handwriting or complex drawing tasks with several thousand points, such as PaHaW t8, Fatigue e06, PECT t08, and PECT t09. The widest mean lengths are observed in PECT, where tasks such as t08 and t09 exceed 8600 points on average, while shorter biometric traces are represented by the signature subsets of MCYT and BiosecurID and by the signature tasks in Fatigue and PECT.

Within individual datasets, task-dependent differences are also marked. EMOTHAW combines relatively short repetitive-pattern tasks with longer drawings and sentence-level handwriting. PaHaW ranges from compact repetitive cursive patterns to longer word-writing tasks and the Archimedean spiral. Fatigue includes very short signatures and spring-like drawings together with much longer repetitive and writing tasks. PECT presents the widest internal span, with short signatures and straight-line traces coexisting with long handwriting and complex drawing tasks. Overall, these results confirm that trajectory length is not a marginal property of the host signal, but one of its main sources of variability.

6.2. In-Air and On-Surface Composition

A second source of variability concerns the relative proportion of in-air and on-surface points. Across all datasets, the pressure signal makes it possible to distinguish between non-contact pen movements ($p = 0$) and points recorded while the pen is on the writing surface ($p > 0$). As shown in the pressure distributions and summarized in Tables 11–16, the balance between these two components changes substantially across datasets and task categories.

Some tasks are almost entirely on-surface. Representative examples include EMOTHAW t4 and t5, PaHaW t1, Fatigue e03, e05, and e07, and PECT t06, all of which exhibit very low in-air percentages. In contrast, other tasks contain a large amount of in-air movement, especially handwriting or drawing activities involving frequent pen lifts and long transitions. High in-air proportions are observed, for example, in EMOTHAW t1, t2, t3, t6, Fatigue e02 and e06, and particularly in PECT t08 and t09, where in-air points account for more than half of the recorded trajectory.

This variability is also visible in the biometric datasets. In both MCYT and BiosecurID, forged signatures exhibit higher in-air proportions than genuine signatures. This indicates that trajectory composition may vary not only with task type, but also with the nature of the signature (genuine or forged). More generally, the results show that online handwriting signals differ not only in total length, but also in how that length is distributed between visible trace and pen-lift movement.

6.3. Implications for Watermark Embedding

From the perspective of watermarking, the descriptive results above suggest that host-signal structure is likely to play a central role in embedding performance. First, longer trajectories provide more samples and therefore a larger potential support for payload. However, raw length alone is not sufficient to characterize embedding conditions, since signals with similar total length may differ substantially in their in-air versus on-surface composition.

Second, the proportion of in-air points is expected to influence the practical behavior of embedding strategies. Signals with a large in-air component may offer more flexibility for embedding information with reduced impact on the visible trace, whereas signals that are almost entirely on-surface may impose stronger constraints on distortion if the geometric appearance of the written trajectory is to be preserved. Conversely, highly regular and predominantly on-surface tasks may provide more stable local structure for some embedding schemes.

Third, task category itself is likely to affect watermarking behavior through its influence on continuity, regularity, and geometric complexity. Signatures, repetitive patterns, extended text, and geometric drawings do not provide the same type of host trajectory, even when their average lengths are similar. Although the experimental evaluation reported in this work is carried out on a reference online signature, the descriptive analysis of datasets and task categories remains useful for placing that signal within a broader structural context.

7. Comparative Results on Payload Capacity

7.1. Payload by Watermarking Method

The results in this section should be interpreted as an analytical capacity illustration rather than as a uniform experimental payload benchmark. The reference case is a hypothetical online handwriting record containing 100 samples in the X coordinate and 100 samples in the Y coordinate, corresponding to 200 coordinate values.

The calculations consider only the X and Y coordinate sequences and follow the embedding support described for each method. For coefficient-based methods, one binary watermark symbol is assumed per modified coefficient. Additional channels, such as pressure, azimuth, or altitude, are not included in this illustration and could provide additional embedding support if they were incorporated into a specific implementation.

For LSB, modifying b bits in each of the 200 coordinate values provides a nominal payload of $200b$ bits. For DCT, using all 100 coefficients from each coordinate sequence and embedding one bit per coefficient provides 200 bits. QIM inserts one bit into each coordinate value and therefore also provides 200 bits.

A one-level Haar decomposition of a 100-sample coordinate sequence produces 50 approximation coefficients. Applying the method separately to X and Y therefore provides 100 approximation coefficients and a nominal DWT payload of 100 bits. For the DFT implementation, a 100-sample real-valued sequence contains 49 positive-frequency coefficients after excluding the DC and Nyquist components. Applying the method separately to both coordinate sequences therefore provides 98 embeddable spectral magnitudes.

The SVD implementation operates on the joint 100×2 coordinate matrix and modifies its two singular values. Its payload is therefore two binary watermark symbols, corresponding to 2 bits. For Difference Expansion, the 100 samples in each coordinate sequence form 50 consecutive pairs. At the maximum evaluated payload rate, $\rho = 1$, one watermark bit is embedded per pair in each coordinate sequence, providing a nominal capacity of 100 bits across X and Y.

Table 21: Analytical nominal payload for a hypothetical online handwriting record containing 100 samples in the X coordinate and 100 samples in the Y coordinate.

Method	Nominal embeddable information in the reference record
LSB (b bits)	$200b$ bits
DCT	200 bits
QIM	200 bits
DWT	100 bits
DFT	98 bits
SVD	2 bits
DE ($\rho = 1.00$)	100 bits

Table 21 shows that the evaluated methods operate under substantially different nominal-capacity regimes. LSB provides the greatest payload flexibility because its capacity increases linearly with the number of modified bits. For example, $b = 3$ provides 600 bits, whereas $b = 5$ provides 1000 bits in the hypothetical reference record.

Under the stated analytical configuration, DCT and QIM each provide 200 bits, DWT provides 100 bits, and DFT provides 98 bits. SVD occupies the opposite extreme, with a payload of only 2 bits because the joint coordinate matrix contains only two singular values.For DE, the nominal payload varies with the payload rate. In the hypothetical reference record, $\rho = 0.5$ corresponds to 50 bits, whereas $\rho = 1$ provides 100 bits.

These values characterize nominal capacity under the common 100-sample-per-axis analytical example. They should not be interpreted as the watermark lengths used to calculate NRMSE, BER, or NC, which were obtained separately using the 4004-point reference signature.

7.2. Payload by Dataset

From the viewpoint of payload capacity, dataset structure is relevant because the available host signal varies substantially across corpora. Table 22 summarizes, for each dataset, the shortest sample, the longest sample, the mean trajectory length, the ratio between the longest and shortest samples, and the mean number of in-air and on-surface points.

This summary is useful because it provides a first estimate of the amount of signal available for watermark embedding before considering any specific method. In particular, the ratio between the longest and shortest samples gives an idea of how heterogeneous the available host space is within each dataset. When this ratio is large, a fixed-length watermark designed for the shortest sample would leave substantial unused space in longer recordings.

An immediate practical implication is that, if the watermark length were chosen using the shortest sample in a dataset as the reference constraint, then the longer samples could accommodate repeated cyclic copies of the same watermark pattern. Such repetition could be exploited as a form of redundancy, potentially improving robustness or increasing resistance to partial degradation in samples with abundant available space.

Table 22: Dataset-level summary of host-signal availability for watermark embedding. Mean in-air and on-surface values are weighted by the number of samples per task or class.

Dataset	Shortest	Longest	Mean	Longest/Shortest	Mean in-air	Mean on-surface
EMOTHAW	143	14398	2942.69	100.69	1196.70	1745.99
PaHaW	405	16071	2285.80	39.68	493.61	1792.18
Fatigue	101	15900	3495.26	157.43	981.36	2513.90
PECT	103	46148	4736.01	448.04	1864.01	2872.00
MCYT	36	5306	581.42	147.39	157.66	423.77
BiosecurID	45	11819	724.45	262.64	200.61	523.85

Table 22 confirms that the datasets differ markedly in their available host-signal space. PECT is the most heterogeneous corpus in this sense, with a longest-to-shortest ratio above 400, whereas PaHaW is much more compact. Fatigue and EMOTHAW also show strong internal variability. These differences suggest that a single fixed payload strategy is unlikely to be equally appropriate across all datasets. Instead, adaptive payload selection or redundancy-aware embedding strategies may be more suitable when watermarking is to be applied over corpora with highly variable sample length.

7.3. Payload by Task Type

Beyond dataset-level differences, host-signal availability also varies according to task type. This is relevant because tasks belonging to the same broad category often share temporal and structural characteristics even when they come from different datasets. The following tables summarize the shortest sample, the longest sample, the mean trajectory length, and the mean number of in-air and on-surface points for each task considered in the taxonomy defined in Section 3.

Table 23: Summary of signature tasks and classes.

Dataset	Task/Class	Shortest	Longest	Mean	Mean in-air	Mean on-surface
Fatigue	e04	101	1568	497.68	141.87	355.81
Fatigue	e08	102	1201	484.59	133.74	350.86
PECT	t07	103	3772	1049.22	266.65	782.58
PECT	t10	170	4359	1111.31	357.94	753.37
MCYT	Genuine	36	2241	444.92	92.52	352.40
MCYT	Forgery	50	5306	717.92	222.80	495.13
BiosecurID	Genuine	45	3625	477.20	104.78	372.43
BiosecurID	Forgery	61	11819	1054.25	328.43	725.82

Table 24: Summary of short-text and word-writing tasks.

Dataset	Task	Shortest	Longest	Mean	Mean in-air	Mean on-surface
EMOTHAW	t3_data	2150	8465	3918.14	1888.78	2029.36
PaHaW	t5	855	7993	2608.48	479.12	2129.36
PaHaW	t6	796	5783	2315.08	496.47	1818.61
PaHaW	t7	585	4423	1469.29	253.91	1215.39
PaHaW	t8	1234	7676	3086.13	1077.48	2008.65
Fatigue	e06	4607	10424	6934.20	3153.10	3781.10
PECT	t08	5210	46148	13180.73	6855.43	6325.30

Table 25: Summary of extended-text tasks.

Dataset	Task	Shortest	Longest	Mean	Mean in-air	Mean on-surface
EMOTHAW	t7_data	1544	8626	3290.97	1261.67	2029.29
Fatigue	e09	2788	9881	4489.23	1791.52	2697.70
PECT	t09	325	42986	8607.94	4378.76	4229.18

Table 26: Summary of geometric-drawing tasks.

Dataset	Task	Shortest	Longest	Mean	Mean in-air	Mean on-surface
EMOTHAW	t1_data	755	11841	2568.58	1123.59	1444.99
EMOTHAW	t2_data	1813	14398	4416.14	2007.17	2408.97
EMOTHAW	t6_data	1369	11541	3811.75	2047.30	1764.45
PaHaW	t1	552	16071	2873.69	63.64	2810.06
Fatigue	e01	1000	8421	2916.36	1096.43	1819.93
Fatigue	e02	2542	13092	5618.97	2489.64	3129.33
Fatigue	e03	1303	5816	2390.51	8.14	2382.37
Fatigue	e07	258	1237	520.89	6.90	513.98
PECT	t01	1162	10004	3549.64	1376.57	2173.08
PECT	t02	3131	25560	7837.26	3826.88	4010.38
PECT	t03	1299	12110	3424.71	608.79	2815.92
PECT	t04	480	3782	1664.91	314.36	1350.55

Table 27: Summary of repetitive-pattern tasks.

Dataset	Task	Shortest	Longest	Mean	Mean in-air	Mean on-surface
EMOTHAW	t4_data	234	3527	1378.93	29.43	1349.50
EMOTHAW	t5_data	143	2421	1202.18	9.90	1192.28
PaHaW	t2	405	4226	1668.01	548.61	1119.40
PaHaW	t3	693	6615	1984.16	487.93	1496.23
PaHaW	t4	614	6827	2305.01	524.55	1780.47
Fatigue	e05	2188	15900	7633.56	18.94	7614.62
PECT	t05	378	5841	1144.35	287.05	857.29
PECT	t06	1312	25193	5927.38	492.84	5434.54

Taken together, these task-wise summaries confirm that payload availability is strongly task-dependent even within the same broad signal modality. Signature tasks are generally among the shortest, whereas short-text, extended-text, and some repetitive-pattern tasks can provide substantially larger host-signal support. Geometric drawings occupy an intermediate but heterogeneous position, depending on their continuity and task design.

8. Distortion Analysis

This section presents a comparative distortion analysis using the reference online signature described in Section 5.2. Because the experiments are performed on a single controlled host signal, the results should be interpreted as relative distortion patterns across watermarking methods and embedding strengths rather than as a complete evaluation across writers, datasets, devices, and task types. Figure 53 shows the complete trajectory, including all recorded points, whereas Figure 54 shows only the samples with non-zero pressure, corresponding to the visible on-surface trace. This distinction is relevant because watermark embedding affects the whole geometric signal, while visual perception depends mainly on the on-surface portion. Visual examples are then provided only for the DCT-based method in order to illustrate how distortion appears on the signature, whereas the quantitative comparison across all methods is reported through the NRMSE results.

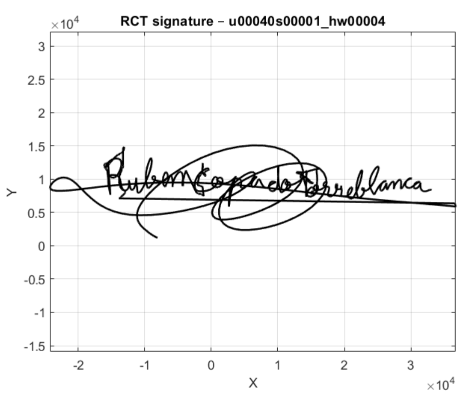


Figure 53: Reference online signature including all recorded points, that is, both on-surface strokes and in-air movements.

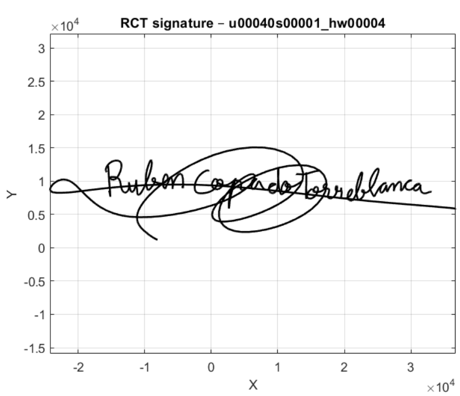


Figure 54: Reference online signature considering only points with non-zero pressure, corresponding to the visible on-surface trace.

8.1. Distortion by Method

The DCT-based scheme is used to illustrate embedding distortion on the reference signature. The cases $\alpha \in \{3, 15, 75\}$ represent low-, medium-, and high-strength embedding. For each case, the analysis includes the watermarked signature, a zoomed view of the letter u, and the geometric difference in the X and Y coordinates. Together, these views show the global trajectory and the local deformation caused by watermarking.

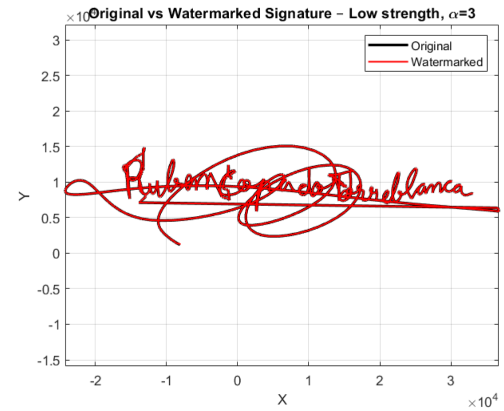


Figure 55: DCT watermarking with $\alpha = 3$: watermarked signature.

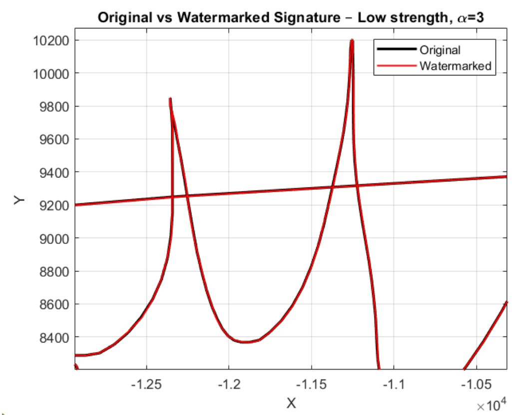


Figure 56: DCT watermarking with $\alpha = 3$: zoomed view of the letter u.

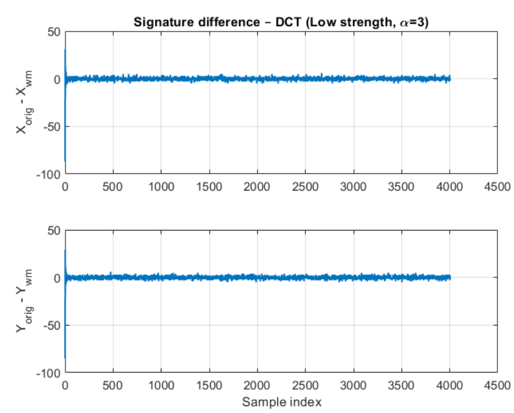


Figure 57: DCT watermarking with $\alpha = 3$: geometric difference between the original and watermarked signatures in the coordinate domain.

At low embedding strength, the watermarked signature remains visually almost indistinguishable from the original. Even in the zoomed view, the deformation is difficult to identify, while the differential representation in Figure 57 reveals that measurable coordinate deviations are already present. This indicates that distortion may remain visually subtle while still affecting the geometric signal.

At intermediate embedding strength, the deformation becomes more apparent. Although the overall signature shape is still preserved, local deviations are easier to observe in both the zoomed trajectory and the coordinate-difference plots.

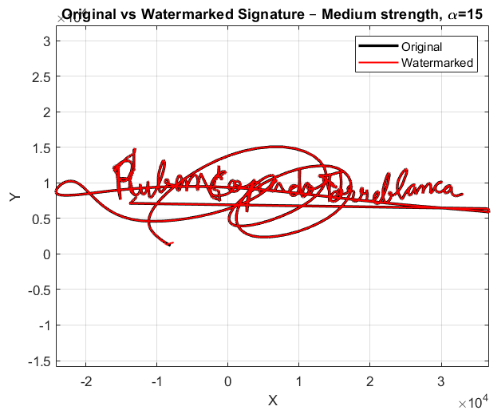


Figure 58: DCT watermarking with α = 15: watermarked signature.

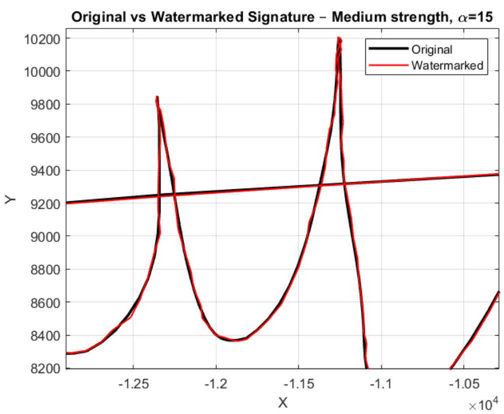


Figure 59: DCT watermarking with α = 15: zoomed view of the letter *u*.

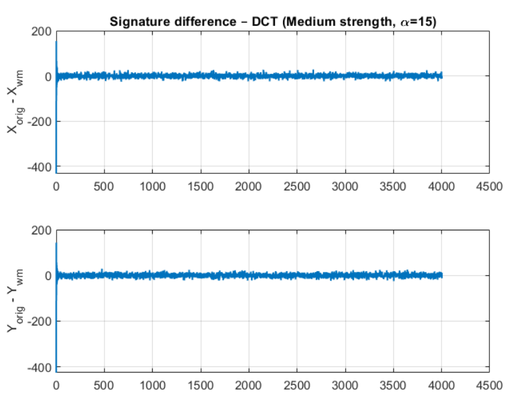


Figure 60: DCT watermarking with α = 15: geometric difference between the original and watermarked signatures in the coordinate domain.

This intermediate case is useful because it illustrates the region in which the watermark begins to have a clearer geometric impact without yet producing the strongest visible degradation. In this sense, it represents a transition between near-imperceptible embedding and clearly noticeable alteration.

At high embedding strength, the distortion becomes much more pronounced. The geometric perturbation is no longer confined to subtle local changes, but affects the trajectory more visibly across the signature.

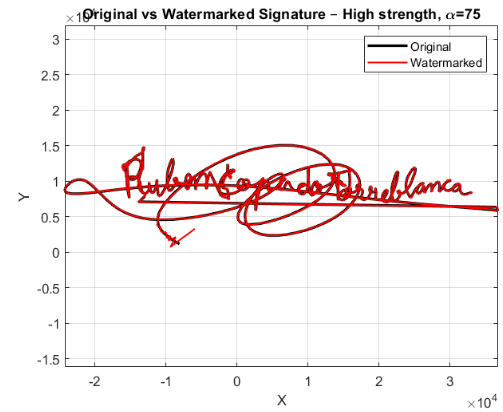


Figure 61: DCT watermarking with α = 75: watermarked signature.

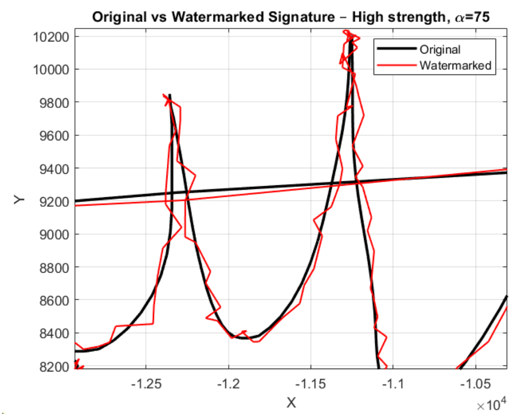


Figure 62: DCT watermarking with α = 75: zoomed view of the letter *u*.

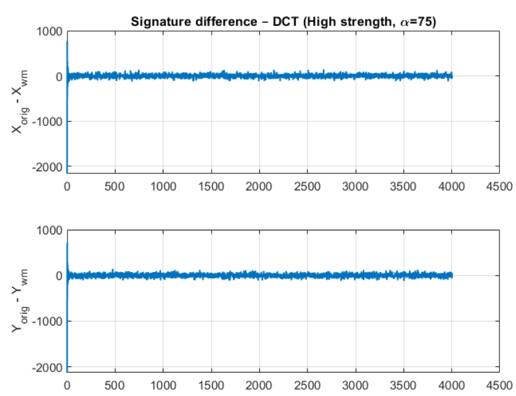


Figure 63: DCT watermarking with α = 75: geometric difference between the original and watermarked signatures in the coordinate domain.

Taken together, Figures 55–63 provide the visual intuition for the rest of the section. At low embedding strength, the signature remains very close to the original and the deformation is mainly revealed by magnification and differential analysis. As α increases, the distortion becomes progressively more evident. This behavior illustrates the central trade-off underlying the distortion analysis: stronger embedding generally improves watermark recoverability, but at the cost of greater alteration of the host signal.

Once this effect has been established visually on the DCT example, the remaining methods are compared quantitatively through the corresponding NRMSE values reported in Table 28. This presentation makes it possible to compare watermarking-induced distortion under a common numerical framework without adding a large number of method-specific figures to the main body of the paper. In all cases, NRMSE provides a compact measure of the overall geometric deviation between the original trajectory and the corresponding watermarked version.

Table 28: NRMSE between the original and watermarked online signature for the compared watermarking methods under low, medium, and high embedding-strength configurations.

Method	Low	Medium	High
DCT ($\alpha = 3, 15, 75$)	4.877×10^{-5}	2.4137×10^{-4}	1.2063×10^{-3}
SVD ($\alpha = 3, 15, 75$)	1.0736×10^{-6}	5.3682×10^{-6}	2.6841×10^{-5}
LSB (1, 3, 5 bits)	1.906×10^{-5}	1.0125×10^{-4}	4.2362×10^{-4}
QIM ($\Delta = 2, 6, 14$)	1.6103×10^{-5}	4.0821×10^{-5}	9.2158×10^{-5}
DWT ($\alpha = 3, 15, 75$)	3.2384×10^{-5}	1.7811×10^{-4}	8.5817×10^{-4}
DE (0.25, 0.5, 1.0 payload)	1.0923×10^{-3}	9.143×10^{-3}	9.3366×10^{-3}
DFT ($\alpha = 3, 15, 75$)	5.0612×10^{-7}	2.5931×10^{-6}	2.0294×10^{-5}

The results confirm that distortion is both method-dependent and strength-dependent. For all methods, stronger embedding leads to higher NRMSE. At the same time, the magnitude of this increase differs markedly across watermarking families. SVD and DFT yield the lowest NRMSE values across the tested configurations, whereas DE produces the largest distortion levels by a wide margin. DCT, DWT, LSB, and QIM occupy intermediate positions, with increasing geometric alteration as embedding strength grows. The DCT example retained in the main text provides a visual illustration of how embedding strength affects the host trajectory, while Table 28 provides the quantitative comparison across all methods.

8.2. Distortion Versus Embedding Strength

For all methods considered in this study, distortion increases with embedding strength. In the transform-domain methods, the control parameter is the embedding gain α; in QIM it is the quantization step Δ; in LSB it is the number of modified least significant bits; and in DE it is the payload rate. Although these parameters are not directly equivalent across methods, in all cases they define progressively stronger embedding conditions.

The DCT visual examples discussed in Section 8 illustrate the effect of embedding strength on geometric distortion, while the NRMSE values reported in Table 28 confirm the same trend quantitatively across the compared methods. Low-strength settings generally produce smaller geometric deviations, whereas medium- and high-strength settings lead to progressively larger distortion levels. Therefore, embedding strength defines a common axis of comparison across watermarking families. Stronger embedding tends to increase the geometric alteration of the host trajectory, even when the specific form and spatial distribution of the distortion depend on the embedding principle.

9. Robustness Analysis

This section presents a controlled robustness analysis using the reference online signature characterized in Section 5.2. The two perturbations defined in Subsection 5.7—additive Gaussian noise and low-pass filtering—were applied under the same attack settings for all watermarking methods. Since the evaluation is restricted to one host signal, the results represent a relative method-level comparison rather than an exhaustive robustness benchmark across writers, datasets, or task categories.

9.1. Robustness Under Gaussian Noise

Gaussian noise provides a direct test of how sensitive each watermarking method is to additive perturbation in the coordinate domain. In the DCT-based case, Figures 64–66 show the difference between the original and extracted watermarks after Gaussian noise for increasing embedding strengths. These figures illustrate the general trend observed throughout the study: weak embedding leads to poor recoverability, whereas stronger embedding significantly improves robustness.

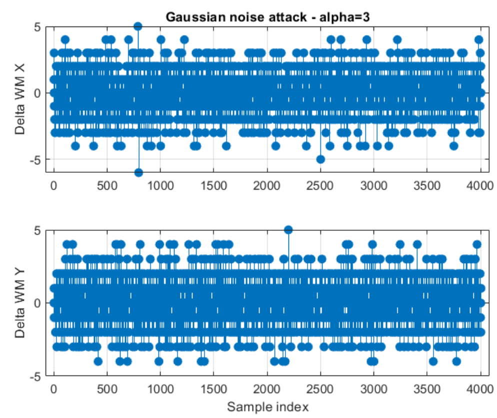


Figure 64: Differences between original and extracted watermarks after Gaussian noise for the DCT-based method with $\alpha = 3$.

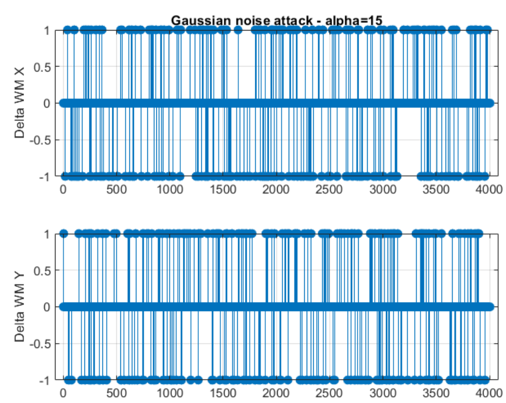


Figure 65: Differences between original and extracted watermarks after Gaussian noise for the DCT-based method with $\alpha = 15$.

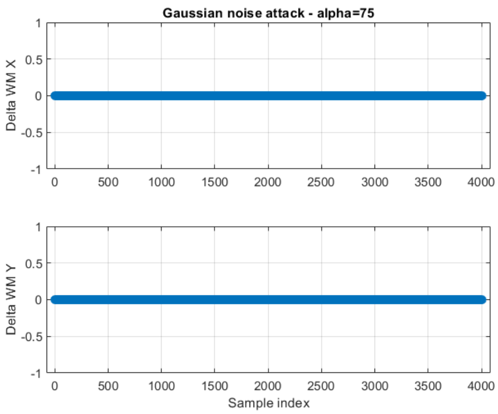


Figure 66: Differences between original and extracted watermarks after Gaussian noise for the DCT-based method with $\alpha = 75$.

The DCT examples shown above provide a visual illustration of watermark degradation under Gaussian noise, while the quantitative robustness results for all methods are summarized in Tables 29 and 30. These tables provide a method-to-method comparison under a common numerical framework, with BER measuring the proportion of incorrectly recovered watermark bits and NC measuring the similarity between the original and extracted watermarks. Reported together, BER and NC provide error-rate and correlation-based views of watermark recovery, while their agreement serves as an internal consistency check.

Table 29: Bit error rate (BER) under additive Gaussian noise for the compared watermarking methods. Results are reported separately for the X and Y coordinate sequences for all methods except SVD. For SVD, the values repeated in the X and Y columns represent the same joint BER computed over the two embedded singular-value watermark symbols and are not coordinate-specific results.

Method	**Low**		**Medium**		**High**	
	BER-X	**BER-Y**	**BER-X**	**BER-Y**	**BER-X**	**BER-Y**
DCT ($\alpha = 3, 15, 75$)	0.35090	0.34441	0.030969	0.036214	0	0
SVD ($\alpha = 3, 15, 75$)	0	0	0	0	0	0
LSB (1, 3, 5 bits)	0.49600	0.49575	0.48834	0.49617	0.41648	0.42657
QIM ($\Delta = 2, 6, 14$)	0.50200	0.48876	0.50325	0.49725	0.36613	0.38162
DWT ($\alpha = 3, 15, 75$)	0.36314	0.33816	0.026474	0.027473	0	0
DE (0.25, 0.5, 1.0 payload)	0.46906	0.49900	0.48452	0.49151	0.49650	0.47652
DFT ($\alpha = 3, 15, 75$)	0.47276	0.47426	0.48576	0.51574	0.40280	0.41479

Table 30: Normalized correlation (NC) under additive Gaussian noise for the compared watermarking methods. Results are reported separately for the X and Y coordinate sequences for all methods except SVD. For SVD, the values repeated in the X and Y columns represent the same joint NC computed over the two embedded singular-value watermark symbols and are not coordinate-specific results.

Method	**Low**		**Medium**		**High**	
	NC-X	**NC-Y**	**NC-X**	**NC-Y**	**NC-X**	**NC-Y**
DCT ($\alpha = 3, 15, 75$)	0.29820	0.31119	0.93806	0.92757	1	1
SVD ($\alpha = 3, 15, 75$)	1	1	1	1	1	1
LSB (1, 3, 5 bits)	0.007992	0.0084915	0.02331	0.007659	0.16703	0.14685
QIM ($\Delta = 2, 6, 14$)	-0.003996	0.022478	-0.0064935	0.0054945	0.26773	0.23676
DWT ($\alpha = 3, 15, 75$)	0.27373	0.32368	0.94705	0.94505	1	1
DE (0.25, 0.5, 1.0 payload)	0.061876	0.001996	0.030969	0.016983	0.006993	0.046953
DFT ($\alpha = 3, 15, 75$)	0.054473	0.051474	0.028486	-0.031484	0.19440	0.17041

The results show a clear separation between fragile and comparatively robust embedding strategies. SVD achieves perfect recovery under Gaussian noise at all tested embedding strengths, with both singular-value watermark symbols recovered correctly (BER = 0, NC = 1). Because only two symbols are embedded, this result should not be interpreted as directly equivalent to perfect recovery of a longer watermark sequence. DCT and DWT also become highly robust at medium and high embedding strengths, with BER values close to zero and NC values close to one for both X and Y. In contrast, LSB, QIM, and DE remain highly sensitive to additive perturbation, with BER values close to random decoding in several configurations and very low NC values. DFT shows limited robustness improvement with increasing embedding strength, but remains clearly less stable than DCT, DWT, and SVD. Overall, the DCT examples presented in the main text, together with the quantitative results in Tables 29 and 30, show that several transform-domain methods benefit more clearly from stronger embedding, whereas spatial, quantization-based, and reversible schemes remain substantially more vulnerable under additive Gaussian noise.

9.2. Robustness Under Low-Pass Filtering

Low-pass filtering evaluates the ability of each method to preserve the watermark under smoothing and attenuation of high-frequency components. The quantitative robustness results for all methods are summarized in Tables 31 and 32. These tables provide a method-to-method comparison under a common numerical framework, with BER measuring the proportion of incorrectly recovered watermark bits and NC measuring the similarity between the original and extracted watermarks. Reported together, BER and NC provide error-rate and correlation-based views of watermark recovery, while their agreement serves as an internal consistency check.

Table 31: Bit error rate (BER) under low-pass filtering for the compared watermarking methods. Results are reported separately for the X and Y coordinate sequences for all methods except SVD. For SVD, the values repeated in the X and Y columns represent the same joint BER computed over the two embedded singular-value watermark symbols and are not coordinate-specific results.

Method	**Low**		**Medium**		**High**	
	BER-X	**BER-Y**	**BER-X**	**BER-Y**	**BER-X**	**BER-Y**
DCT ($\alpha = 3, 15, 75$)	0.38761	0.035465	0.20854	0	0	0
SVD ($\alpha = 3, 15, 75$)	1	1	1	1	0	0
LSB (1, 3, 5 bits)	0.030969	0.022228	0.019564	0.015068	0.042008	0.046104
QIM ($\Delta = 2, 6, 14$)	0.03047	0.022478	0.003996	0.0034965	0.00074925	0.0014985
DWT ($\alpha = 3, 15, 75$)	0.0054945	0.0044955	0.0014985	0.0004995	0.000999	0
DE (0.25, 0.5, 1.0 payload)	0.091816	0.15170	0.11289	0.18681	0.073926	0.088412
DFT ($\alpha = 3, 15, 75$)	0.50225	0.51324	0.50375	0.50975	0.47526	0.14693

Table 32: Normalized correlation (NC) under low-pass filtering for the compared watermarking methods. Results are reported separately for the X and Y coordinate sequences for all methods except SVD. For SVD, the values repeated in the X and Y columns represent the same joint NC computed over the two embedded singular-value watermark symbols and are not coordinate-specific results.

Method	**Low**		**Medium**		**High**	
	NC-X	**NC-Y**	**NC-X**	**NC-Y**	**NC-X**	**NC-Y**
DCT ($\alpha = 3, 15, 75$)	0.22478	0.92907	0.58292	1	1	1
SVD ($\alpha = 3, 15, 75$)	-1	-1	-1	-1	1	1
LSB (1, 3, 5 bits)	0.93806	0.95554	0.96087	0.96986	0.91598	0.90779
QIM ($\Delta = 2, 6, 14$)	0.93906	0.95504	0.99201	0.99301	0.9985	0.997
DWT ($\alpha = 3, 15, 75$)	0.98901	0.99101	0.997	0.999	0.998	1
DE (0.25, 0.5, 1.0 payload)	0.81637	0.69661	0.77423	0.62637	0.85215	0.82318
DFT ($\alpha = 3, 15, 75$)	-0.0044978	-0.026487	-0.0074963	-0.01949	0.049475	0.70615

The results show a robustness profile under low-pass filtering that differs in some respects from the Gaussian-noise case. DWT provides the most favorable overall performance, achieving near-perfect recovery across all tested configurations, with very low BER and NC values close to one for both coordinate sequences. DCT also improves substantially with embedding strength and reaches perfect recovery at high strength, although its low- and medium-strength behavior is less balanced between the X and Y coordinates. QIM shows a strong and consistent response under smoothing, with BER decreasing and NC approaching one as the quantization step increases. LSB remains comparatively stable under this attack, with low BER and high NC despite its weaker behavior under Gaussian noise. DE shows moderate robustness, with noticeable degradation but still substantially better performance than under Gaussian perturbation. By contrast, DFT performs poorly under low-pass filtering, with BER values close to random decoding at low and medium strengths and negative NC values in those configurations, especially for the X coordinate. SVD exhibits a strongly discrete response under low-pass filtering. At low and medium embedding strengths, both singular-value watermark symbols are recovered with reversed polarity, yielding BER = 1 and NC = −1. At high embedding strength, both symbols are recovered correctly, yielding BER = 0 and NC = 1. Taken together, the quantitative results in Tables 31 and 32 indicate that even low-pass filtering can substantially affect watermark recovery, although the relative behavior of the methods differs from that observed under additive Gaussian noise.

9.3. Robustness Versus Embedding Strength

The effect of embedding strength on robustness is method-dependent and attack-dependent rather than uniform across all watermarking families. In DCT and DWT, increasing the embedding strength clearly improves recoverability under both Gaussian noise and low-pass filtering, leading to very low BER values and NC values close to one in the strongest configurations. In QIM, increasing Δ also improves extraction stability, particularly under low-pass filtering, where the method shows consistently strong performance. In LSB, the effect of increasing the number of modified bits is limited and does not produce the same degree of robustness improvement, especially under Gaussian noise. In DE, increasing payload does not translate into a comparable robustness gain, and the method remains relatively sensitive to perturbation across configurations. SVD exhibits strongly attack-dependent behavior: both singular-value watermark symbols are recovered correctly under Gaussian noise at all tested strengths. Under low-pass filtering, SVD fails at low and medium strengths, with both embedded symbols extracted with inverted polarity, and recovers both symbols correctly only at the highest configuration. DFT shows only modest improvement with stronger embedding and remains comparatively fragile, especially under low-pass filtering.

These trends are supported by the quantitative results reported in Tables 29–32. Low-strength configurations are often insufficient for reliable extraction, but the extent to which stronger embedding improves recovery differs markedly across methods. In some cases, such as DCT, DWT, and QIM, stronger embedding leads to substantial gains in robustness. In others, such as LSB, DE, and DFT, the improvement is limited or inconsistent. SVD is particularly noteworthy because it combines very low payload with perfect robustness under Gaussian noise, yet its behavior under low-pass filtering reveals a much stronger dependence on embedding threshold.

Therefore, embedding strength acts as an important control axis across watermarking methods, but its practical effect depends strongly on both the embedding family and the attack model. Stronger embedding is not equally effective in all methods, and this difference is central to the comparative interpretation of robustness in online handwriting watermarking.

Table 33 condenses the distortion and robustness results obtained under the strongest evaluated configuration of each method.

Table 33: Summary of distortion and robustness under the strongest evaluated configuration of each method. BER and NC are averaged across the X and Y coordinate sequences, except for SVD, for which they refer jointly to the two embedded watermark bits.

Method	**NRMSE**	**Gaussian BER / NC**	**Low-pass BER / NC**
LSB ($b = 5$)	4.2362×10^{-4}	0.422/0.157	0.044/0.912
DE ($\rho = 1.00$)	9.3366×10^{-3}	0.487/0.027	0.081/0.838
QIM ($\Delta = 14$)	9.2158×10^{-5}	0.374/0.252	0.001/0.998
DCT ($\alpha = 75$)	1.2063×10^{-3}	0/1	0/1
DWT ($\alpha = 75$)	8.5817×10^{-4}	0/1	< 0.001/0.999
DFT ($\alpha = 75$)	2.0294×10^{-5}	0.409/0.182	0.311/0.378
SVD ($\alpha = 75$)	2.6841×10^{-5}	0/1	0/1

9.4. Robustness–Payload Trade-Off

The robustness results show that embedding capacity and resistance to perturbation are not directly aligned. LSB provides the highest nominal payload flexibility, but its watermark recovery remains poor under Gaussian noise. DE offers reversibility, although it combines comparatively high distortion with limited robustness under the tested attacks. QIM occupies an intermediate position, with recovery improving as the quantization step increases, particularly under low-pass filtering.

Among the transform-domain methods, DCT and DWT provide the most balanced behavior in the evaluated reference setting, combining moderate nominal capacity with strong recovery at medium and high embedding strengths. SVD represents a different operating point, since it provides extremely low payload but perfect recovery under Gaussian noise and strongly threshold-dependent behavior under low-pass filtering. DFT produces very low geometric distortion, but this does not translate into comparable robustness.

Overall, the results obtained for the evaluated reference signature indicate that payload, distortion, and robustness should be considered jointly. Under the tested configurations, methods with greater nominal capacity were not necessarily more resistant to perturbation, while methods showing strong recovery sometimes operated under substantial capacity restrictions. The practical implications of these observed trade-offs are discussed in Section 10.7.

10. Discussion

The main outcome of this study is not that one watermarking method dominates the others in all scenarios, but that the compared methods occupy different regions of the payload–distortion–robustness space. This interpretation is informed by the cross-dataset characterization of host-signal structure and by the controlled experimental comparison carried out on a reference online signature. For this reason, the discussion is organized around the three main comparative axes of the study—capacity, signal preservation, and robustness—together with two contextual dimensions that are especially relevant in online handwriting, namely the role of in-air trajectory and the differences between datasets collected in e-health and e-security scenarios.

10.1. New Insights From the Structured Comparison

The comparative results obtained in the evaluated setting indicate that watermarking behavior cannot be adequately characterized using a single metric or the embedding domain alone. Three observations support this conclusion. First, nominal capacity and robustness are not aligned: LSB provides the largest payload flexibility but remains highly vulnerable to Gaussian perturbation, whereas SVD provides extremely limited capacity but perfect recovery under the same attack in the evaluated setting. Second, low geometric distortion does not necessarily imply reliable watermark recovery. DFT produces some of the lowest NRMSE values but remains comparatively fragile, particularly under low-pass filtering. Third, robustness is strongly attack-dependent: methods such as LSB, QIM, and SVD exhibit substantially different behavior under Gaussian noise and smoothing.

The cross-dataset characterization adds a complementary host-signal perspective. Trajectory length determines the nominal amount of available embedding support, while the proportion of in-air and on-surface points determines how that support is distributed between visible and non-visible trajectory segments. These descriptors therefore provide a basis for adaptive watermarking policies in which payload or embedding location is selected according to the structure of the host signal. However, because the present experiments do not independently evaluate these policies across datasets, this interpretation should be understood as a structured hypothesis derived from the descriptive evidence rather than as a demonstrated causal relationship.

The observed method-specific behavior can also be interpreted in relation to the corresponding embedding and extraction mechanisms. LSB modifies the least significant components of the coordinate values directly; consequently, even small coordinate perturbations can alter the embedded information, which is consistent with its poor recovery under Gaussian noise. In DCT and DWT, increasing α enlarges the modification introduced in the selected transform coefficients, making the embedded information easier to distinguish after attack, although at the cost of increased geometric distortion. The DFT results provide a particularly relevant counterexample: its very low NRMSE does not result in comparable robustness, suggesting that limited host distortion alone is not an adequate indicator of watermark stability under filtering and coordinate rounding.

The QIM results indicate that increasing the separation between quantization regions through a larger Δ can improve recovery under smoothing. The SVD results should be interpreted more cautiously because the watermark contains only two values. Consequently, its BER and NC measurements are much more discrete than those obtained from longer watermark sequences, and perfect recovery is not directly equivalent to perfect recovery of a high-payload watermark. Furthermore, the specific implementations evaluated in this study use different amounts and types of reference information during watermark recovery. Therefore, the observed robustness differences reflect not only the embedding domain, but also the watermark length, the operating configuration, and the recovery procedure implemented for each method. These implementation-specific requirements should not be interpreted as universal properties of the corresponding watermarking families.

10.2. Which Methods Maximize Payload?

From the viewpoint of nominal embedding capacity, the methods considered in this work do not operate under the same regime. The reference comparison reported in Table 21 shows that LSB provides the greatest payload flexibility, since its capacity increases linearly with the number of modified least significant bits. Under the common reference case of 100 samples in X and 100 samples in Y, LSB provides 200 bits for $b = 1$, 600 bits for $b = 3$, and 1000 bits for $b = 5$. Its capacity therefore exceeds that of the other evaluated methods when more than one bit per coordinate value is modified.

Under the analytical configuration reported in Table 21, DCT and QIM each provide 200 bits, DWT provides 100 bits, and DFT provides 98 bits. SVD provides only 2 bits. DE also has a configuration-dependent nominal payload, which varies with the selected payload rate. All payload values in this comparison refer exclusively to the hypothetical 100-sample-per-axis reference record.

Therefore, nominal payload should be interpreted jointly with distortion, robustness, and reversibility rather than as an isolated optimization objective.

10.3. Which Methods Better Preserve the Signal?

The distortion analysis shows that signal preservation depends not only on the amount of embedded information, but also on the embedding strategy. The DCT example used as a visual introduction makes clear that low-strength embedding may leave the watermark almost imperceptible in the written trace, whereas stronger configurations produce progressively more visible deformation. The NRMSE values reported in Table 28 also show that the magnitude of distortion differs substantially across methods.

At a quantitative level, SVD and DFT produce the lowest NRMSE values, whereas DE produces the largest distortion. DCT, DWT, LSB, and QIM occupy intermediate positions. These differences indicate that signal preservation is strongly method-dependent and should be interpreted together with payload and robustness.

Accordingly, in the analyzed reference setting, geometric distortion was implementation-dependent rather than consistently determined by the embedding family. DFT and SVD produced the lowest reported NRMSE values, QIM and LSB remained below DCT and DWT, and DE produced the largest distortion. These comparisons should be interpreted together with the different payloads, parameter scales, and embedding supports of the methods.

10.4. Which Methods Are More Robust?

The robustness analysis under Gaussian noise and low-pass filtering shows that the relative behavior of the compared methods is not identical across perturbation types, although some broad patterns remain clear. Direct spatial embedding tends to be more fragile under Gaussian perturbation, whereas several transform-domain methods benefit more clearly from stronger embedding. In particular, DCT and DWT exhibit consistently favorable robustness behavior as embedding strength increases, reaching near-perfect recovery in their strongest configurations under both attacks. QIM occupies an intermediate but controllable position, with robustness improving substantially as the quantization step increases, especially under low-pass filtering. By contrast, DE remains comparatively sensitive to perturbation despite its reversibility. LSB also shows attack-dependent behavior: it remains fragile under Gaussian noise, but performs much better under low-pass filtering than under additive perturbation. DFT does not exhibit the same robustness level as DCT and DWT, particularly under low-pass filtering, where its recovery quality remains limited. SVD is especially sensitive to the type of attack: it achieves perfect robustness under Gaussian noise at all tested strengths, but under low-pass filtering it fails at low and medium strengths and only recovers at the highest configuration.

Overall, the robustness results show that embedding strength is beneficial mainly when the embedding mechanism places the watermark in components that remain stable under the corresponding attack. Consequently, robustness rankings should always be reported together with the attack model and the evaluated parameter range.

10.5. Role of In-Air Trajectory in Embedding Capacity

One of the most relevant observations arising from the dataset characterization is that online handwriting signals differ not only in total length, but also in how that length is distributed between in-air and on-surface movement. This distinction is especially important in watermarking because in-air trajectory may act as an additional host-signal region whose geometric relevance is not identical to that of the visible written trace.

In datasets and tasks with large in-air proportions, the available host signal is substantially expanded beyond the visible pen trace. From a capacity perspective, this means that two signals with similar on-surface writing content may nevertheless offer very different embedding opportunities if one of them includes longer pen-lift and transition trajectories. In this sense, in-air movement can be interpreted as a structurally relevant source of additional payload support.

At the same time, the present study does not yet isolate in-air embedding as an independent experimental condition. Therefore, the role of in-air trajectory should be interpreted here as a strong structural hypothesis supported by the dataset analysis, rather than as a directly quantified causal result. Even so, the descriptive evidence suggests that exploiting in-air segments more explicitly could become an important strategy for increasing payload while controlling visible distortion in future watermarking designs.

10.6. Differences Between e-Health and e-Security Datasets

The datasets considered in this work were collected under different application contexts, and this difference is reflected in their signal structure. The e-security corpora, represented mainly by MCYT and BiosecurID, are centered on signature acquisition and biometric variability under genuine and forgery conditions. Their signals are relatively compact when compared with many of the handwriting and drawing tasks present in the other datasets, although clear class-related differences appear, especially in the higher in-air proportions observed in forged signatures.

By contrast, the datasets more closely associated with e-health or general handwriting analysis include a much wider range of task categories, such as geometric drawings, repetitive patterns, isolated words, and extended text. This increases the span of available trajectory lengths and the diversity of in-air versus on-surface compositions. From the perspective of watermarking, this broader variability makes them particularly useful for understanding how host-signal structure conditions payload and embedding behavior.

These differences are important for positioning the study. The e-security datasets are especially relevant when watermarking is interpreted in a biometric context, whereas the broader handwriting datasets define heterogeneous host-signal conditions under which future method-level evaluations should be conducted. Rather than treating these two groups as competing scenarios, the present work benefits from their complementarity: one group contributes application realism, and the other contributes structural variability.

10.7. Practical Recommendations by Application

The comparative results indicate that watermarking method selection should be driven by the operational requirements of the target application rather than by a single global ranking. The main design criteria are the amount of information to be embedded, the need to recover the original trajectory, the expected signal processing operations, and the acceptable level of geometric alteration.

When high payload is the main objective, LSB provides the most direct and scalable solution because its capacity increases with the number of modified bits. This option may be suitable for controlled storage or transmission scenarios in which the signal is not expected to undergo coordinate perturbations. However, its limited robustness under Gaussian noise makes it less appropriate for persistent ownership or traceability marks that must survive signal modification.

When exact recovery of the original host signal is required, Difference Expansion is the most relevant method among those evaluated. This property may be useful in archival, biomedical, or forensic scenarios in which the original coordinate trajectory must be restored after data extraction. Nevertheless, reversibility is guaranteed only in the absence of attacks, and the comparatively high distortion and limited robustness observed in the experiments restrict its suitability for hostile or strongly processed environments.

For applications requiring persistent identifiers, such as traceability or source association, DCT and DWT provide the most favorable overall robustness–distortion compromise in the evaluated setting. QIM may be appropriate when a tunable compromise is required, since the quantization step provides direct control over the balance between distortion and recovery, particularly under smoothing. SVD is suitable only when a very compact identifier is sufficient, because its strong recovery is associated with an extremely limited payload. The DFT implementation produced low geometric distortion but insufficient robustness to support its use as a persistent watermark under the evaluated attacks.

Host-signal structure should also be considered when selecting the payload. Longer trajectories provide greater nominal embedding support, while signals with substantial in-air movement may offer additional embedding opportunities outside the visible trace. In practical systems, this suggests using adaptive payloads based on trajectory length and evaluating whether information should be distributed differently between in-air and on-surface segments.

These recommendations are restricted to the evaluated reference signature, parameter configurations, and attack conditions. Before deployment, the selected method should be validated on the target device, handwriting task, writer population, and expected processing chain. In biometric applications, its effect on recognition performance should also be assessed, since low geometric distortion does not by itself guarantee preservation of biometric utility.

11. Limitations

The limitations of this study are organized according to three complementary dimensions: experimental design and method comparability, dataset heterogeneity, and application scope. The most important limitation is the asymmetry between the breadth of the cross-dataset characterization and the narrower scope of the controlled experiments. While multiple datasets and task categories are analyzed from a structural perspective, distortion and robustness are experimentally evaluated on a single reference signature. Consequently, the reported method rankings should be interpreted as controlled reference-level evidence rather than as population-level or cross-dataset performance estimates.

The remaining limitations concern the comparability of the watermarking configurations, the restricted set of attacks, differences between dataset acquisition protocols, and the absence of biometric-performance evaluation. These limitations are detailed below in methodological, dataset-related, and scope-specific categories.

11.1. Methodological Limitations

The first limitation of the present study is that the watermarking methods do not share the same intrinsic capacity model. Although a common reference case was defined to facilitate comparison, the embedding support and capacity constraints differ across methods. Under the specific implementation assumptions described in Section 7.1, the nominal payload can be calculated analytically for all seven methods. However, these values are configuration-dependent and should not be interpreted as directly equivalent capacity measures. In addition, the payload of transform-domain methods depends on the adopted decomposition and coefficient-selection strategy.

A second limitation concerns the nature of the robustness evaluation. Only two perturbations were considered, namely additive Gaussian noise and low-pass filtering. These attacks are useful for illustrating comparative behavior, but they do not exhaust the range of transformations that may affect online handwriting signals in real applications. More aggressive filtering, geometric transformations, resampling, compression-like operations, or mixed attacks could lead to different robustness profiles.

A third methodological limitation concerns the use of the single reference signature described in Section 5.2. This design reduces host-signal variability and enables a controlled comparison across methods, but it does not capture differences between writers, tasks, acquisition devices, or datasets. Therefore, the distortion and robustness results should not be interpreted as population-level performance estimates. The broader dataset analysis provides structural context for the host signals, but additional multidataset experiments would be required to determine the generalizability of the observed method rankings.

The use of a fixed reference signature was intentional because it isolates method-level differences under identical host-signal conditions. The resulting evidence therefore supports a controlled comparison but does not establish cross-writer or cross-dataset generalization. Broader validation across signals, writers, devices, and task categories is consequently included as a specific direction for future work.

11.2. Dataset-Related Limitations

The datasets included in this work are heterogeneous in acquisition protocol, task design, and application context. This heterogeneity is valuable for structural analysis, but it also limits strict comparability. Differences in recording conditions, task instructions, and corpus objectives mean that dataset-level contrasts should be interpreted cautiously and not as if all corpora had been collected under a unified benchmark protocol.

A further limitation concerns the availability of contextual metadata. In particular, the PECT dataset contributes substantial signal variability and a rich task inventory, but its acquisition context is less fully documented in the present manuscript than that of some other corpora. This does not prevent its use for signal characterization, but it does limit the interpretability of some application-level comparisons.

11.3. Scope Limitations

The present paper focuses on payload, distortion, and robustness, but does not evaluate the impact of watermark embedding on biometric verification performance. In particular, no analysis is provided here of how embedding affects measures such as FAR, FRR, EER, or DET behavior. This is an important limitation because, in biometric applications, preserving watermark information is not sufficient if the embedded signal loses recognition utility.

In the same way, the study does not experimentally separate embedding in in-air and on-surface segments, even though the dataset analysis suggests that this distinction may be highly relevant. As a result, the role of in-air trajectory is discussed as a structurally motivated interpretation rather than as an experimentally isolated effect.

More broadly, the paper should be understood as a comparative study of watermarking behavior in online handwriting signals, not as a full end-to-end biometric deployment study. Its contribution lies in clarifying the interaction between host-signal structure and watermarking trade-offs, while leaving application-specific optimization and recognition-level validation for future work.

Overall, the use of a single reference signature is the limitation with the greatest effect on the generalizability of the experimental findings. The method-specific payload definitions and the use of only two attack conditions also limit the interpretation of the comparison as a universal ranking. Dataset heterogeneity and the absence of biometric verification experiments mainly affect the transfer of the results to specific practical applications. Accordingly, the conclusions are restricted to the evaluated signal, parameter configurations, implementation choices, and perturbation settings.

12. Conclusion

The literature review identified a lack of structured comparative evidence connecting host-signal characteristics, payload capacity, geometric distortion, reversibility, and robustness in online handwriting watermarking. This study addresses that gap by combining a cross-dataset characterization of trajectory length, task type, and in-air/on-surface composition with a controlled comparison of seven representative watermarking methods using common distortion and robustness metrics.

The results show that the available embedding support depends strongly on the structure of the host trajectory and that no watermarking method simultaneously maximizes payload, imperceptibility, reversibility, and robustness. Transform-domain methods, particularly DCT and DWT, provided the most favorable robustness behavior under the evaluated conditions, whereas LSB offered greater payload flexibility but limited resistance to perturbations. QIM provided a tunable compromise, Difference Expansion enabled reversibility in attack-free conditions, and SVD achieved strong recovery for a very limited payload.

Overall, the study partially closes the identified research gap by providing a common framework that places method-level watermarking performance alongside cross-dataset host-signal characterization and by making the principal trade-offs directly comparable. However, the distortion and robustness experiments were conducted using a single reference signature and a restricted set of parameters and attacks. Consequently, the reported method rankings should be interpreted as controlled reference-level evidence rather than as a universal cross-dataset benchmark.

13. Future Work

Future work will extend the present study through a set of specific experimental stages.

First, the controlled evaluation will be expanded to multiple writers, datasets, and task categories. The seven watermarking methods will be applied to representative signatures, isolated words, extended texts, geometric drawings, and repetitive patterns. Multiple watermark realizations will be evaluated for each host signal, reporting mean values and variability for NRMSE, BER, and NC. This extension will determine whether the method rankings observed for the current reference signature remain consistent under different host-signal conditions.

Second, separate embedding policies will be evaluated for in-air and on-surface trajectory segments. These experiments will compare embedding over the complete signal, only in-air samples, only on-surface samples, and adaptive combinations of both regions. Comparisons will be performed under matched payload conditions to determine whether in-air embedding can increase capacity while reducing visible geometric distortion.

Third, robustness evaluation will be extended beyond the two attack configurations used in the present study. Gaussian noise will be examined over several values of σ, and low-pass filtering will be tested using different cutoff frequencies and filter orders. Additional perturbations, including coordinate resampling, scaling, translation, rotation, temporal modification, and combined attacks, will also be considered. These experiments will make it possible to construct robustness curves rather than relying on a single severity level for each attack.

Fourth, watermarking performance will be connected to biometric utility. Experiments on signature datasets such as MCYT and BiosecurID will evaluate whether watermark embedding and subsequent attacks affect verification performance in terms of FAR, FRR, EER, and DET curves. This analysis will determine whether low geometric distortion is sufficient to preserve the discriminative information required by biometric systems.

Finally, future comparisons will include matched operating conditions based on equal payload or equal distortion. This will complement the present within-method low-, medium-, and high-strength configurations and provide a more standardized basis for comparing fundamentally different watermarking families. The resulting evidence may support adaptive watermarking strategies in which the embedding method, payload, and strength are automatically selected according to trajectory length, task category, in-air proportion, and application requirements.

Funding This work was partly supported by project PID2023-146644OB-I00, funded by MICIU/AEI (10.13039/501100011033) and by the European Union through the FEDER program.

Author Contributions Ruben Copado-Torreblanca and Marcos Faundez-Zanuy contributed equally to the conceptualization, methodology, investigation, analysis, writing, and revision of this work. Both authors read and approved the final manuscript.

Conflicts of Interest The authors declare no conflicts of interest.

Use of AI Tools Generative AI tools, including ChatGPT (OpenAI), were used to assist with LaTeX drafting and formatting, to generate and refine manuscript text based on the authors' explanations, to transform author-provided computational outputs into tables, and to identify potential improvements in clarity and presentation. These tools were also used to assist in the generation, revision, and debugging of MATLAB and Python code. All code was reviewed, executed, tested, and validated by the authors, and all numerical results and AI-assisted content were independently verified by them. The authors take full responsibility for the accuracy, originality, and integrity of the manuscript.

References

[1] M. Faundez-Zanuy, "Signature recognition state-of-the-art," IEEE Aerospace and Electronic Systems Magazine, **20**(7), 28–32, 2005, doi:10.1109/MAES.2005.1499249.

[2] A. Martin, G. Doddington, T. Kamm, M. Ordowski, M. Przybocki, "The DET curve in assessment of detection task performance," in Proc. 5th European Conference on Speech Communication and Technology (Eurospeech 1997), 1895–1898, 1997, doi:10.21437/Eurospeech.1997-504.

[3] M. Faundez-Zanuy, "On-line signature recognition based on VQ-DTW," Pattern Recognition, **40**(3), 981–992, 2007, doi:10.1016/j.patcog.2006.06.007.

[4] M. Faundez-Zanuy, M. Diaz, "On the Use of First and Second Derivative Approximations for Biometric Online Signature Recognition," in I. Rojas, G. Joya, A. Catala, editors, Advances in Computational Intelligence, 461–472, Springer Nature Switzerland, Cham, 2023.

[5] M. R. Nilchiyan, R. B. Yusof, S. E. Alavi, "Statistical Online Signature Verification Using Rotation-Invariant Dynamic Descriptors," in Proceedings of the Asian Control Conference (ASCC), 1–6, Kota Kinabalu, Malaysia, 2015.

[6] A. Chadha, D. Jyoti, M. M. Roja, "Rotation, Scaling and Translation Analysis of Biometric Signature Templates," International Journal of Computer Technology and Applications, **2**(5), 1419–1425, 2011.

[7] M. Gomez-Barrero, J. Galbally, J. Fierrez, J. Ortega-García, R. Plamondon, "Enhanced On-Line Signature Verification Based on Skilled Forgery Detection Using Sigma-LogNormal Features," in Proceedings of the International Conference on Biometrics (ICB), 2015.

[8] M. Faundez-Zanuy, "Analysis of Inclinations in Genuine Signatures and Skilled Forgeries," in Proceedings of the 22nd Conference of the International Graphonomics Society (IGS 2025), Polytechnique Montréal, Canada, 2025, https://www.igs2025.org.

[9] F. A. P. Petitcolas, R. J. Anderson, M. G. Kuhn, "Information Hiding – A Survey," Proceedings of the IEEE, **87**(7), 1062–1078, 1999, doi:10.1109/5.771065.

[10] M. Faundez-Zanuy, "Online Signature Watermarking in the Transform Domain," Cognitive Computation, **17**(2), 79, 2025, doi:10.1007/s12559-025-10436-y.

[11] M. Faundez-Zanuy, "Blind Watermarking of Online Handwritten Signals Based on DCT," in Proceedings of the 33rd European Signal Processing Conference (EUSIPCO), 1337–1341, Isola delle Femmine – Palermo, Italy, 2025, iSBN: 978-9-46-459362-4.

[12] M. Faundez-Zanuy, "Comprehensive Analysis of Least Significant Bit and Difference Expansion Watermarking Algorithms for Online Signature Signals," Expert Systems with Applications, **267**, 126214, 2025, doi:10.1016/j.eswa.2024.126214.

[13] M. Faundez-Zanuy, M. Diaz, M. A. Ferrer, "Fragile Watermarking in the Pressure Signal of Online Signatures for Integrity Checking," IEEE Access, **14**, 60489–60500, 2026, doi:10.1109/ACCESS.2026.3684266.

[14] A. Hayat, S. S. Ali, V. Kumar, A. K. Bhateja, "SigTem: A Non-Invertible Technique for Online Signature Template Protection," Expert Systems with Applications, **283**, 127646, 2025, doi:10.1016/j.eswa.2025.127646.

[15] A. K. Gupta, C. Chakraborty, B. Gupta, "Watermarking of EEG Data to Provide Security Based on DWT-SVD," International Journal of Distributed Systems and Technologies, 2022, doi:10.4018/IJDST.307902.

[16] P. M. Tripathi, "Watermarking of ECG Signals Compressed Using Fourier Decomposition Method," Multimedia Tools and Applications, 2022, doi:10.1007/s11042-021-11492-w.

[17] J. Rani, A. Anand, "SecECG: Secure Data Hiding Approach for ECG Signals," Multimedia Tools and Applications, **83**(14), 42885–42905, 2024, doi:10.1007/s11042-023-17049-3.

[18] Y. Naderahmadian, "ECG Signal Watermarking Using QR Decomposition," Physical and Engineering Sciences in Medicine, 2024, doi:10.1007/s13246-024-01480-3.

[19] P. Andreev, A. Denisova, V. Fedoseev, "Reversible Watermarking for Electrocardiogram Protection," Sensors, **25**(7), 2185, 2025, doi:10.3390/s25072185.

[20] S. Gull, S. A. Parah, "Advances in medical image watermarking: a state of the art review," Multimedia Tools and Applications, **83**(1), 1407–1447, 2024, doi:10.1007/s11042-023-15396-9.

[21] A. Mehto, N. Mehra, "Techniques of Digital Image Watermarking: A Review," International Journal of Computer Applications, **128**(9), 2015, doi:10.5120/ijca2015906629.

[22] M. Begum, M. S. Uddin, "Digital Image Watermarking Techniques: A Review," Information, **11**(2), 110, 2020, doi:10.3390/info11020110.

[23] M. Choudhary, S. Sharma, "A Comprehensive Review of Digital Watermarking Methods," International Journal of Engineering Technology and Computing Research, 2024.

[24] S. Gaur, V. Barthwal, "An Extensive Analysis of Digital Image Watermarking Techniques," International Journal of Intelligent Systems and Applications in Engineering, 2024.

[25] J. Varghese, O. Bin Hussain, S. Subash, T. Abdul Razak, "An Effective Digital Image Watermarking Scheme Incorporating DCT, DFT and SVD Transformations," PeerJ Computer Science, 2023.

[26] A. Benoraira, K. Benmahammed, N. Boucenna, "Blind Image Watermarking Technique Based on Differential Embedding in DWT and DCT Domains," EURASIP Journal on Advances in Signal Processing, 2015, doi:10.1186/s13634-015-0239-5.

[27] W. Alomoush, O. A. Khashan, A. Alrosan, H. H. Attar, A. Almomani, F. Alhosban, S. N. Makhadmeh, "Digital Image Watermarking Using Discrete Cosine Transformation Based Linear Modulation," Journal of Cloud Computing, 2023.

[28] N. I. Yassin, N. M. Salem, M. I. El-Adawy, "QIM Blind Video Watermarking Scheme Based on Wavelet Transform and Principal Component Analysis," Alexandria Engineering Journal, **53**(4), 833–843, 2014.

[29] S. R. Moulick, S. Arora, C. Jain, P. K. Panigrahi, "Reliable SVD Based Semi-Blind and Invisible Watermarking Schemes," arXiv preprint arXiv:1503.01934, 2015.

[30] N. Chawla, V. Singh, "A Review of DWT and PCA Based Digital Watermarking Schemes," International Journal of Advanced Research in Computer Science, 2018.

[31] L. Likforman-Sulem, A. Esposito, M. Faundez-Zanuy, S. Clemencon, G. Cordasco, "EMOTHAW: A Novel Database for Emotional State Recognition from Handwriting and Drawing," IEEE Transactions on Human-Machine Systems, **47**(2), 273–284, 2017, doi:10.1109/THMS.2016.2635441.

[32] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, M. Faundez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease," Artificial Intelligence in Medicine, **67**, 39–46, 2016, doi:10.1016/j.artmed.2016.01.004, epub 2016 Feb 4.

[33] M. Faundez-Zanuy, J. Lopez-Xarbau, M. Diaz, M. Garnacho-Castaño, "On the Analysis of Saturated Pressure to Detect Fatigue," in A. Parziale, M. Diaz, F. Melo, editors, Graphonomics in Human Body Movement. Bridging Research and Practice from Motor Control to Handwriting Analysis and Recognition, volume 14285 of *Lecture Notes in Computer Science*, 47–57, Springer, Cham, 2023, doi:10.1007/978-3-031-45461-5_4.

[34] F. Candela, S. Romeo, M. Faundez-Zanuy, P. Ferrer-Ramos, "Cognitive Impairment Detection Based on Frontal Camera Scene While Performing Handwriting Tasks," Cognitive Computation, **16**(3), 1004–1021, 2024, doi:10.1007/s12559-024-10279-z.

[35] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J.-J. Igarza, C. Vivaracho, D. Escudero, Q.-I. Moro, "MCYT baseline corpus: a bimodal biometric database," IEE Proceedings - Vision, Image and Signal Processing, **150**(6), 395–401, 2003, doi:10.1049/ip-vis:20031078.

[36] J. Fierrez, J. Galbally, J. Ortega-Garcia, M. R. Freire, F. Alonso-Fernandez, D. Ramos, D. T. Toledano, J. Gonzalez-Rodriguez, J. A. Siguenza, J. Garrido-Salas, E. Anguiano, G. Gonzalez-de Rivera, R. Ribalda, M. Faundez-Zanuy, J. A. Ortega, V. Cardeñoso-Payo, A. Viloria, C. E. Vivaracho, Q. I. Moro, J. J. Igarza, J. Sanchez, I. Hernaez, C. Orrite-Uruñuela, F. Martinez-Contreras, J. J. Gracia-Roche, "BiosecurID: A Multimodal Biometric Database," Pattern Analysis and Applications, **13**(2), 235–246, 2010, doi:10.1007/s10044-009-0151-4.

[37] J. Tian, "Reversible Data Embedding Using a Difference Expansion," IEEE Transactions on Circuits and Systems for Video Technology, **13**(8), 890–896, 2003, doi:10.1109/TCSVT.2003.815962.

[38] B. Chen, G. W. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," IEEE Transactions on Information Theory, **47**(4), 1423–1443, 2001, doi:10.1109/18.923725.

[39] C.-T. Hsu, J.-L. Wu, "Multiresolution Watermarking for Digital Images," IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, **45**(8), 1097–1101, 1998, doi:10.1109/82.718818.

[40] V. Solachidis, I. Pitas, "Circularly Symmetric Watermark Embedding in 2-D DFT Domain," IEEE Transactions on Image Processing, **10**(11), 1741–1753, 2001, doi:10.1109/83.967401.

[41] R. Liu, T. Tan, "An SVD-Based Watermarking Scheme for Protecting Rightful Ownership," IEEE Transactions on Multimedia, **4**(1), 121–128, 2002, doi:10.1109/6046.985560.