IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

IEEE CONSUMER TECHNOLOGY SOCIETY SECTION

Received March 10, 2022, accepted April 5, 2022, date of publication April 14, 2022, date of current version April 21, 2022.

*Digital Object Identifier 10.1109/ACCESS.2022.3167413*

# Online Signature Verification Using Locally Weighted Dynamic Time Warping via Multiple Fusion Strategies

**MANABU OKAWA, (Member, IEEE)**
Forensic Science Laboratory, Metropolitan Police Department, Tokyo 100-8929, Japan
e-mail: m_okawa.lab1@m.ieice.org

This work was supported by JSPS KAKENHI under Grant JP21H04330 and Grant JP22H04430.

**ABSTRACT** Online signature verification is the process of using a dynamic signature verification system to confirm the writer's identity. It can be used as a security system to confirm entrance applications and password substitutes, as well as a forensic tool to assist an expert's investigation. This study proposes a novel online signature verification system based on a single-template strategy to improve performance in real-world scenarios. It employs discriminative mean signature template sets as well as fusion strategies of multiple local weighting and warping schemes, for dynamic time warping (DTW). The first step is to generate a set of user-specific mean signature templates for each feature using a recent time-series averaging method, namely, Euclidean barycenter-based DTW barycenter averaging. Then, using multiple and direct matching points between the mean signature templates and references for dependent and independent DTW, we obtain a local weighting estimate considering local stability sequences. Furthermore, we develop fusion strategies for calculating locally weighted DTW sets and concatenating them as a feature vector for each warping, followed by the construction of a support vector machine (SVM) classifier. Finally, in the verification phase, we use the single-template technique to compute a discriminative fused score using SVMs between the mean signature template sets and a query sample. The effectiveness of the proposed method is demonstrated by extensive experimental results obtained using three public online signature datasets: SVC2004 Task1/Task2 and MCYT-100.

**INDEX TERMS** Biometrics, forensics, signature verification, time series analysis, fusion strategy, dynamic time warping, support vector machine.

## I. INTRODUCTION

Handwriting is a common means of communication in our daily lives, and signatures are socially and legally accepted as a form of individual authentication based on each person's behavioral characteristics and unique features. Because of the widespread adoption of consumer electronics applications and products (e.g., tablets, phablets, and cell/mobile phones), online signature verification systems have recently been used in biometrics [1]–[3] and forensics [4]–[7].

To extract a number of features from dynamic signatures, both parametric and functional approaches are used. The former method depicts signatures as a set of parameters or vectors (e.g., total signature duration, number of pen ups/downs, and average/maximum speed), whereas the latter method represents signatures as time functions (e.g., pen position trajectory, pressure, and velocity). The functional technique has consistently outperformed the parametric approach [1]–[3]. Consequently, the functional approach is the focus of this study.

The useful method has been adopted for template matching using a distance measure such as dynamic time warping (DTW). Template matching can be classified into two types: multiple-template and single-template strategies. The former compares the distances between a query pattern and each reference using certain statistical measures (e.g., max,

The associate editor coordinating the review of this manuscript and approving it for publication was Yoonsik Choe.

median, min, and mean), whereas the latter computes a single template chosen or generated from a reference set. The single-template strategy has many advantages, including speed, security, and tolerance based entirely on one-to-one matching. However, its accuracy is lower than that of the multiple-template approach [8].

A novel single-template strategy based on a time-series averaging method, namely, Euclidean barycenter-based DTW barycenter averaging (EB-DBA) [8], and locally/globally weighted DTW (LG-DTW), has recently been proposed [9], [10]. The study [9] used multiple matching points (MMPs) [11] for the local weighting estimate and the variable importance obtained using gradient boosting (GB) [12] for the global weighting estimate, whereas the other study [10] used direct matching points (DMPs) [13] rather than MMPs for the local weighting estimate in LG-DTW. Based on the recent demand for high-speed systems in the Big Data era, these single-template strategies can reduce calculation complexity while achieving efficient performance, which results in a higher level of accuracy.

However, there are certain issues with verification performance in real-world applications that are listed below:

(1) Using MMPs and DMPs separately diminishes their complementary effects and may result in the loss of detailed local stability information that existed between mean signature templates and reference sets.

(2) The use of MMPs and DMPs individually makes it difficult for the system to adapt to changes in writing conditions (e.g., device and signal types), template aging, and skilled forgery attacks, all of which frequently occur in real-world scenarios.

Consequently, we obtained the following solutions for these challenges:

(1) We introduced a modified local weighting scheme for DTW using both MMPs and DMPs (namely, LM-DTW and LD-DTW) for dependent and independent warping to incorporate more detailed and flexible local stability information and effectively minimize intra-class discrepancies.

(2) To enhance inter-user variability, we used the multiple fusion strategies: the representation-level fusion to concatenate LM-DTW and LD-DTW as a single vector (namely, F-DTW) for each warping, followed by the score-level fusion to combine each score from multiple support vector machine (SVM) classifiers [14] built for each warping.

Note that this study is an extension of our previous research [15]. Then, as detailed below, we revised the previous method and performed additional experiments:

(1) We updated the global weighting scheme using GB in the score-level fusion with SVM to improve F-DTW discriminative power and performance.

(2) In addition to the previous experiments using an SVC2004 Task1 dataset [16], we conducted extensive experiments using two public datasets, SVC2004 Task2 [16] and MCYT-100 [17], to confirm the generalization performance of the proposed method.

The remainder of the paper is organized as follows: In Section II, we review recent online signature verification methods and distance measures relevant to this study. Throughout Section III, we present the proposed online signature verification method. In Section IV, we explain the experimental methods and results. In the penultimate part (Section V), we discuss the findings and their applicability in real-world scenarios. Finally, in Section VI, we present the conclusion.

## II. RELATED WORK

### A. ONLINE SIGNATURE VERIFICATION

In the past decade, numerous online signature verification systems have been proposed [1]–[3]. The system can be divided in two types of matching methods: model-based and distance-based approaches. Model-based approaches describe data distribution by employing generative models (e.g., Gaussian models [18] and hidden Markov models (HMMs) [19]) and discriminative models (e.g., SVMs [20], convolutional neural networks (CNNs) [21], and recurrent neural networks (RNNs) [22]). Distance-based approaches use distance measures such as DTW [23] to match query signatures with reference sets. Distance-based approaches are superior in forensic situations with limited data availability for enrollment because a model-based approach would suffer from overfitting issues.

Among multiple distance-based systems, template matching is commonly used for online signature verification [18]. Template matching approaches include single-template and multiple-template strategies. The single-template strategy has additional benefits, such as speed, security, and tolerance, all of which are in high demand in today's digital era. It does not, however, outperform as well as the multiple-template strategy [18].

To address the abovementioned challenges, a recent study [8] proposed an effective single-template strategy that uses mean signature templates created using a novel time-series averaging method known as EB-DBA. The template creation method, as described in [9]–[11], [13], and [24], increased the possibility of template matching in online signature verification.

Distance measures are extensively used in template matching techniques to calculate the dissimilarity between the templates and query signatures of varying lengths.

### B. DISTANCE MEASURES

Lockstep and elastic distance measures are the two types of distance measures [25]. Lockstep measures are calculated by strictly aligning the time-series indices using one-to-one mappings, such as Euclidean distances. These measures, however, are susceptible to noise, outliers, and basic shape variations with irregular lengths. To address these limitations, elastic measures such as DTW [23] have been proposed for optimally aligning time-series indices with a one-to-many mapping based on dynamic programming.

#### 1) DTW

We can calculate DTW for $K$-dimensional multivariate time series using independent and dependent warping [26]. For each time sequence, a DTW with independent warping ($\mathrm{DTW_I}$) is calculated, assuming that each DTW is a distance measure with a one-dimensional (1D) trajectory in 1D Euclidean space. The DTW with dependent warping ($\mathrm{DTW_D}$) is derived directly as a single DTW corresponding to the set of time sequences, assuming the considered $K$-dimensional time series as a 1D trajectory in $K$-dimensional Euclidean space. Recent online signature verification studies [8]–[10], [24] demonstrate that $\mathrm{DTW_I}$ and $\mathrm{DTW_D}$ have distinct/complementary discriminative powers. The DTW calculation is described in detail below.

Assuming $\boldsymbol{A}$ and $\boldsymbol{B}$ are two $K$-dimensional multivariate time series of different lengths, $I$ and $J$, respectively, they are defined as follows:

$$\boldsymbol{A} = \{A_k\}_{k=1}^{K} = \{\boldsymbol{a}(1), \boldsymbol{a}(2), \ldots, \boldsymbol{a}(i), \ldots, \boldsymbol{a}(I)\},$$
$$\boldsymbol{B} = \{B_k\}_{k=1}^{K} = \{\boldsymbol{b}(1), \boldsymbol{b}(2), \ldots, \boldsymbol{b}(j), \ldots, \boldsymbol{b}(J)\},$$

where $\boldsymbol{a}(i) = \{a_k(i)\}_{k=1}^{K}$ and $\boldsymbol{b}(j) = \{b_k(j)\}_{k=1}^{K}$, respectively.

Then, $\mathrm{DTW_I}$ and $\mathrm{DTW_D}$ can be computed as follows.

##### a: DTW WITH INDEPENDENT WARPING ($\mathrm{DTW_I}$)

First, the $I \times J$ cost matrix is developed using the cost function $d(\cdot, \cdot)$ between two time points defined as follows:

$$d(a_k(i), b_k(j)) = (a_k(i) - b_k(j))^2. \tag{1}$$

Then, a warping path $W = \{w_z\}_{z=1}^{Z}$ with $\max(I, J) \le Z \le (I + J - 1)$ is derived based on the cost matrix, satisfying the boundary, continuity, and monotonicity conditions set forth in [23].

Finally, $k$th dimensional $\mathrm{DTW_I^k}$ can be defined as follows:

$$\mathrm{DTW_I^k} = DTW(A_k, B_k) = \min_{W} \left\{ \sum_{z=1}^{Z} d(w_z) \right\}, \tag{2}$$

where $d(w_z) = d(a_k(i), b_k(j))$ corresponds to $i$ and $j$ at position $z$ in the warping path by recursively calculating the cumulative distance as follows:

$$D(i,j) = d(a_k(i), b_k(j)) + \min \begin{cases} D(i, j-1), \\ D(i-1, j-1), \\ D(i-1, j). \end{cases} \tag{3}$$

##### b: DTW WITH DEPENDENT WARPING ($\mathrm{DTW_D}$)

In a manner similar to $\mathrm{DTW_I}$, $\mathrm{DTW_D}$ can be defined by calculating it with dependent warping to obtain a single distance from the set of time sequences as follows:

$$\mathrm{DTW_D} = DTW(\boldsymbol{A}, \boldsymbol{B}), \tag{4}$$

where $d(\cdot, \cdot)$ in Eq. (1) is replaced with

$$d(\boldsymbol{a}(i), \boldsymbol{b}(j)) = \sum_{k=1}^{K} (a_k(i) - b_k(j))^2. \tag{5}$$

Consequently, DTW can find the best alignment and try to minimize distances between time sequences of varying lengths; thus, it has been extensively used in online signature verification [2], [18].

However, DTW is sensitive to noise and outliers in time sequences because it needs to pair all elements of a sequence. Certain weighting methods for DTW have been proposed to compensate for such limitations.

#### 2) WEIGHTING SCHEMES FOR DTW

There are two types of DTW weighting schemes: local and global weighting schemes. The specifics are outlined below.

##### a: LOCAL WEIGHTING SCHEME

The local weighing scheme includes a weighting function that adds weight to the DTW cost function between matching points.

A previous study [27] proposed a weighted DTW (WDTW), which includes a multiplicative weight penalty based on the distances between the the warping path points. Using this method, the cost matrix is updated to incorporate a modified logistic weight function that adds weight to the DTW cost function between the reference and test points. A sliding window DTW (SW-DTW) approach has been proposed as a viable option [28]. SW-DTW used the modified DTW cost function with a window function to account for context by incorporating a weighted average of the neighboring distances.

Certain studies used DTW matching with a weighting scheme to estimate local stability domains in online signature verification. The study [29] proposed a stability-modulated DTW (SM-DTW) to incorporate the most similar parts of two signatures in the DTW distance measure. Other studies [30], [31] used a weighted DMP to incorporate information from the DTW matching to determine the most stable domains for each signer.

These studies highlight the possibility of using local weighting methods for DTW. Most of the previous approaches, however, required a multiple-template strategy and/or adequate parameter optimization, resulting in high computational complexity. To address these issues, we proposed the locally weighted DTW [11], [13], in which we calculate the local stability of the mean signature template set using MMPs or DMPs and apply weights to the DTW cost functions.

##### b: GLOBAL WEIGHTING SCHEME

The global weighting scheme includes feature weighing/selection and applies to DTW or its variants. Canonical time warping (CTW) [32] combines DTW with canonical correlation analysis (CCA) to compute spatial projections

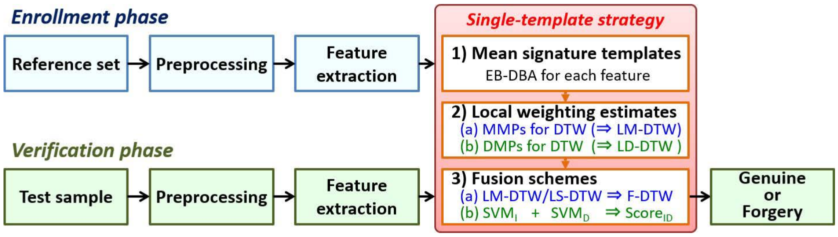



**FIGURE 1.** Outline of the proposed online signature verification method.

and identify the linear combinations of variables between two different multivariate sequences. CTW can incorporate feature weighting/selection as well as a dimensionality reduction mechanism to align signals of varying dimensions. Another study [24] proposed a novel single-template strategy that employs a global weighting scheme to combine multiple DTWs while weighting them with variable importance via GB.

These studies bring to light the potential for weighting methods to be applied to DTW. The separate/independent use of local or global weighting, however, results in DTW's limited discriminative power. Consequently, we proposed a novel single-template strategy to address these issues.

## III. PROPOSED METHOD

### A. OUTLINE

The proposed online signature verification method is shown in Fig. 1.

We used preprocessing after online signature input to improve quality and extract common function-based features. During the enrollment phase, we used a single template strategy, which included mean signature template generation based on the EB-DBA and DTW fusion strategies based on multiple local weighting and warping schemes using the reference set. During the verification phase, we calculated the fused score between a test sample and a purported user's mean signature templates. Finally, the system determines whether the test sample is genuine or forged based on whether the fuss score is less or greater than a user-specified threshold.

### B. PREPROCESSING

To address the natural fluctuations in signatures because of writing conditions, we used the following normalization approach for horizontal and vertical pen coordinates $\{x(t), y(t)\}$ in [8] and [13]:

$$\hat{x}(t) = \frac{x(t) - x_g}{x_{max} - x_{min}}, \quad \hat{y}(t) = \frac{y(t) - y_g}{y_{max} - y_{min}} \tag{6}$$

where $(x_g, y_g)$ is the centroid of a signature; $\{x_{min}, y_{min}\}$ and $\{x_{max}, y_{max}\}$ are the minimum and maximum values of $\{x(t), y(t)\}$ for $t = 1, 2, \ldots, T$, respectively.

In the following, the pen coordinates $\{x(t), y(t)\}$ show the preprocessed functions, $\{\hat{x}(t), \hat{y}(t)\}$.

### C. FEATURE EXTRACTION

This study employs seven common function-based features [13], [18], [19], which are then normalization using z-score.

- There are three unique features: horizontal and vertical pen coordinates $x(t)$, $y(t)$, and pen pressure $p(t)$.
- Path-tangent angle $\theta(t)$, path velocity magnitude $v(t)$, log curvature radius $\rho(t)$, and total acceleration magnitude $\alpha(t)$ are four additional features derived from the original $x(t)$, $y(t)$ as follows:

$$\theta(t) = \arctan(\dot{y}(t)/\dot{x}(t)), \tag{7}$$

$$v(t) = \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}, \tag{8}$$

$$\rho(t) = \log\left(v(t)/\dot{\theta}(t)\right), \tag{9}$$

$$\alpha(t) = \sqrt{\dot{v}(t)^2 + (v(t) \cdot \dot{\theta}(t))^2}, \tag{10}$$

where the derivatives of discrete-time signals are computed using a second-order regression that removes small noisy variations using the following formula [33]:

$$\dot{f}(t) = \frac{\sum_{\epsilon=1}^{2} \epsilon(f(t+\epsilon) - f(t-\epsilon))}{2\sum_{\epsilon=1}^{2} \epsilon^2}. \tag{11}$$

Note that certain digital devices provide original signals without the use of pen pressure (e.g., the SVC2004 Task1 used in this study). In that case, we choose six of the above seven function-based features that do not include $p(t)$.

### D. SINGLE-TEMPLATE STRATEGY

The proposed single-template strategy is broken down into three steps: 1) the mean signature templates, 2) local weighting estimates, and 3) fusion schemes (Fig. 2). The details of each step are described in the subsections that follow with references to the definitions in Section II-B1.

#### 1) MEAN SIGNATURE TEMPLATES

The single-template strategy employs user-specific mean signature templates (i.e., multiple prototypes corresponding to each user's feature) via EB-DBA [8] to account for intra-user variations across all reference samples.

EB-DBA is a time-series averaging method that iteratively refines the Euclidean barycenter (EB) sequence to minimize its DTW to average target sequences using an expectation–maximization scheme. In particular, we developed an EB sequence of $N$ references in which the elements

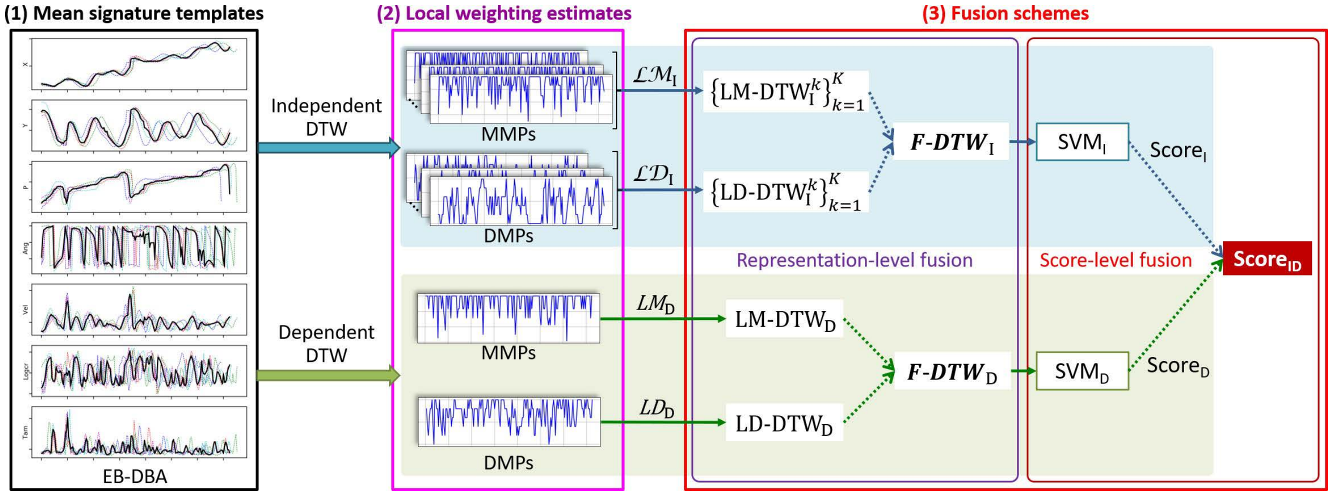



**FIGURE 2.** Process of the proposed single-template strategy: (1) mean signature template set creation per feature (pen coordinates "X" and "Y," pen pressure "P," path-tangent angle "Ang," path velocity magnitude "Vel," log curvature radius "Logcr," and total acceleration magnitude "Tam") through EB-DBA (solid black lines) using the five original reference sequences (dashed lines in various colors); (2) local weighting estimates with MMPs and DMPs for independent and dependent DTW; (3) fusion schemes in representation-level and score-level fusions through the SVM models. In the verification phase, we finally obtain the $\mathrm{Score}_{\mathrm{ID}}$.

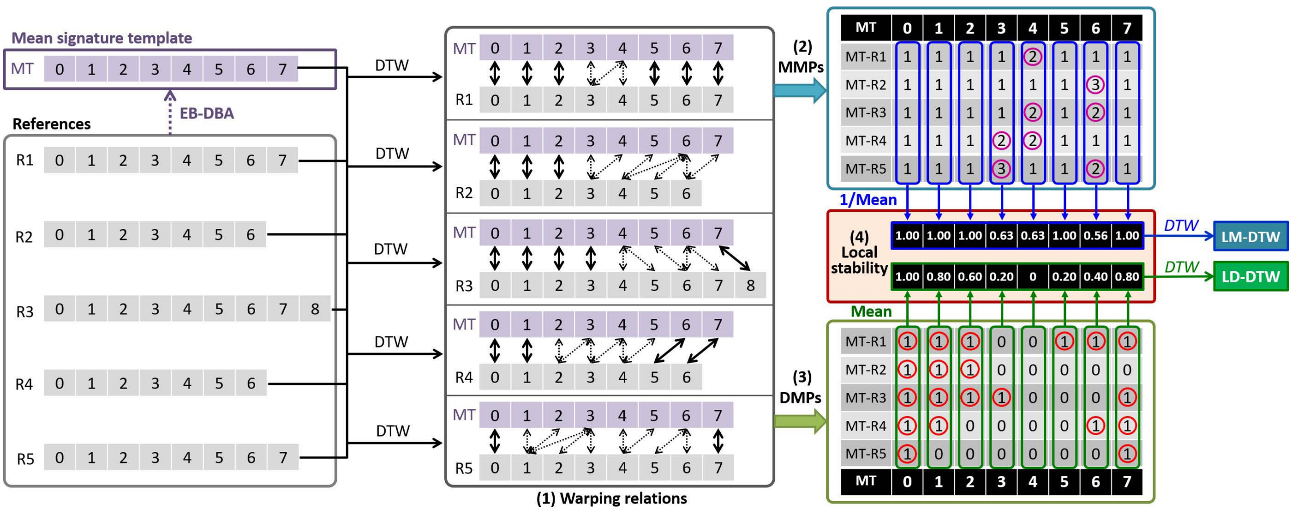



**FIGURE 3.** Example of the estimation process for local stability using toy samples: (1) the warping relations between the mean signature template "MT" and five references, "R1" to "R5," of varying lengths ("0" to "8" in each sequence revealing the indices), (2) the MMPs, (3) the DMPs, and (4) the local stability sequences, which are applied to DTW calculations to obtain LM-DTW/LD-DTW.

are resampled using linear interpolation to reach their average length $I$ equally. We then peformed the DBA [34] from the original $N$ references using the EB sequence as the initial sequence.

We obtained $K$ mean signature templates of lengths $I$ for each user using EB-DBA from $K$ function-based features (i.e., $K = 6$ for the SVC2004 Task1; $K = 7$ for the SVC2004 Task2 and the MCYT-100 in this study).

### 2) LOCAL WEIGHTING ESTIMATES

To determine local fluctuations and incorporate intra-user variations into the distance measure, we estimate local stability regions in signatures. MMPs and DMPs were used in this study to estimate the complementary local stability of mean signature templates (Fig. 3).

#### *a: MMPs*

MMPs [11] detect multiple matching points in DTW trajectories where the mean signature template set and the references are distorted significantly. Consequently, the MMP sequence indicates the mean signature template sequence's local instability, and the inverse of the averaged MMP sequence $\{mmp_i\}_{i=1}^{I}$ can be considered as the local stability.

For simplicity, let us assume that mean signature template $A$ corresponds to an $I$-length univariate time sequence and that the original set of $N$ references $\mathcal{B} = \{B^n\}_{n=1}^{N}$ corresponds to a $J_n$-length univariate time sequence. The estimation process of MMP-based local stability can then be summarized as follows.

(1) We first calculate the standard DTW for each warping between $A$ and $\mathcal{B}$ and obtain a set of $N$ optimal warping paths as per the following formula:

$$\mathcal{W}(A, \mathcal{B}) = \{W^n(A, B^n)\}_{n=1}^{N}.$$

(2) Next, we compute $N$ MMP sequences from $\mathcal{W}(A, \mathcal{B})$ and obtain the averaged MMP sequence:

$$\{mmp_i\}_{i=1}^{I} = \left\{\frac{1}{N}\sum_{n=1}^{N} c_i^n\right\}_{i=1}^{I} \tag{12}$$

where $c_i^n$ is the cardinality of a set, represented as $card\{\cdot\}$, belonging to the $i$th point of $A$ defined as follows:

$$c_i^n = card\{(i^n, j^n) \in W^n(A, B^n) \mid i^n = i\},$$

By following the steps outlined above, we finally obtained $I$-length local weight sequences $\mathcal{LM}_{\mathrm{I}} = \{LM_{\mathrm{I}}^k\}_{k=1}^{K}$ for independent warping and $LM_{\mathrm{D}}$ for dependent warping, defined as follows:

$$\begin{aligned} LM_{\mathrm{I}}^k &= \{lm_{\mathrm{I}}^k(1), lm_{\mathrm{I}}^k(2), \ldots, lm_{\mathrm{I}}^k(i), \ldots, lm_{\mathrm{I}}^k(I)\} \\ LM_{\mathrm{D}} &= \{lm_{\mathrm{D}}(1), lm_{\mathrm{D}}(2), \ldots, lm_{\mathrm{D}}(i), \ldots, lm_{\mathrm{D}}(I)\} \end{aligned}$$

where $lm(i) = 1/mmp_i$ is $0 < lm(i) \leq 1$; $lm(i)$ is 1 when all matching point pairs are DMPs and approaches 0 with an increase in the number of MMPs.

#### b: DMPs

DMPs [13] detect averaged matching points in DTW trajectories where one-to-one matching relations exist between the mean signature template set and all references.

The following describes the estimation process for DMP-based local stability.

(1) Similar to the MMPs, we begin by calculating a set of $N$ optimal warping paths

$$\mathcal{W}(A, \mathcal{B}) = \{W^n(A, B^n)\}_{n=1}^{N},$$

where $W^n(A, B^n) = \{(p_z^n, q_z^n)\}_{z=1}^{Z_n}$ with $1 \leq p_z^n \leq I$, $1 \leq q_z^n \leq J_n$, and $\max(I, J_n) \leq Z_n \leq I + J_n - 1$ describes a $Z_n$-length warping path.

(2) Then, using the set of warping paths, we compute the $N$ DMP sequences. When the multiplicity of the warping relation for each component is defined as the number of consecutive occurrences of the component index in $W^n(A, B^n)$, the multiplicities corresponding to the respective matching components of $p_z^n$ and $q_z^n$ are as follows:

$$\begin{aligned} m_i^n &= card\{(p_z^n, q_z^n) \in W^n(\boldsymbol{A}, \boldsymbol{B}^n) \mid p_z^n = i\}, \\ m_j^n &= card\{(p_z^n, q_z^n) \in W^n(\boldsymbol{A}, \boldsymbol{B}^n) \mid q_z^n = j\}. \end{aligned}$$

The $i$th point of $A$, where the multiplicity simultaneously satisfies both $m_i^n = 1$ and $m_j^n = 1$, can be defined as a DMP.

By following the steps outlined above, we finally obtained $I$-length local weight sequences $\mathcal{LD}_{\mathrm{I}} = \{LD_{\mathrm{I}}^k\}_{k=1}^{K}$ for independent warping and $LD_{\mathrm{D}}$ for dependent warping, defined as follows:

$$\begin{aligned} LD_{\mathrm{I}}^k &= \{ld_{\mathrm{I}}^k(1), ld_{\mathrm{I}}^k(2), \ldots, ld_{\mathrm{I}}^k(i), \ldots, ld_{\mathrm{I}}^k(I)\}, \\ LD_{\mathrm{D}} &= \{ld_{\mathrm{D}}(1), ld_{\mathrm{D}}(2), \ldots, ld_{\mathrm{D}}(i), \ldots, ld_{\mathrm{D}}(I)\}, \end{aligned}$$

where

$$\{ld(i)\}_{i=1}^{I} = \left\{\frac{1}{N}\sum_{n=1}^{N} c_i^n\right\}_{i=1}^{I} \tag{13}$$

with

$$c_i^n = \begin{cases} 1, & \text{if an } i\text{th point is a DMP} \\ 0, & \text{otherwise} \end{cases}$$

and $0 \leq ld(i) \leq 1$; $ld(i)$ is 0 and 1 when all matching point pairs are MMPs and DMPs, respectively.

### 3) FUSION SCHEMES

After calculating $\mathcal{LM}_{\mathrm{I}}$ and $LM_{\mathrm{D}}$ with MMPs and $\mathcal{LD}_{\mathrm{I}}$ and $LD_{\mathrm{D}}$ with DMPs, we used fusion schemes to obtain a discriminative score while maximizing the inter-user variations. This study employed two fusion strategies [5]: representation-level and score-level fusions.

#### a: REPRESENTATION-LEVEL FUSION

For representation-level fusion, we first obtained a locally weighted DTW with MMPs and DMPs (i.e., LM-DTW and LD-DTW, respectively), and then combined them in a single vector, F-DTW.

To obtain LM-DTW, the cost function $d(\cdot, \cdot)$ between the two points of the considered time series, as defined in Eqs. (1) and (5), can be rewritten by weighting them by the corresponding local weight sequences, $\mathcal{LM}_{\mathrm{I}}$ and $LM_{\mathrm{D}}$, as follows:

$$d(a_k(i), b_k(j)) = lm_{\mathrm{I}}^k(i) \times (a_k(i) - b_k(j))^2, \tag{14}$$

$$d(\boldsymbol{a}(i), \boldsymbol{b}(j)) = lm_{\mathrm{D}}(i) \times \sum_{k=1}^{K} (a_k(i) - b_k(j))^2. \tag{15}$$

Similar to LM-DTW, to obtain LD-DTW, the cost function $d(\cdot, \cdot)$ in Eqs. (1) and (5) can be rewritten as follows using $\mathcal{LD}_{\mathrm{I}}$ and $LD_{\mathrm{D}}$:

$$d(a_k(i), b_k(j)) = ld_{\mathrm{I}}^k(i) \times (a_k(i) - b_k(j))^2, \tag{16}$$

$$d(\boldsymbol{a}(i), \boldsymbol{b}(j)) = ld_{\mathrm{D}}(i) \times \sum_{k=1}^{K} (a_k(i) - b_k(j))^2. \tag{17}$$

Consequently, we obtained $\{\text{LM-DTW}_{\mathrm{I}}^k\}_{k=1}^{K}$ and $\{\text{LD-DTW}_{\mathrm{I}}^k\}_{k=1}^{K}$ for independent warping, and LM-DTW$_{\mathrm{D}}$ and LD-DTW$_{\mathrm{D}}$ for dependent warping.

Finally, we obtained F-DTW with independent and dependent warping (i.e., ***F-DTW***$_{\mathrm{I}}$ and ***F-DTW***$_{\mathrm{D}}$, respectively), described below:

$$
\begin{aligned}
\boldsymbol{F\text{-}DTW}_{\mathrm{I}} &= \{\text{F-DTW}_{\mathrm{I}}^{u}\}_{u=1}^{2K} \\
&= (\text{LM-DTW}_{\mathrm{I}}^{1}, \ldots, \text{LM-DTW}_{\mathrm{I}}^{K}, \\
&\quad \text{LD-DTW}_{\mathrm{I}}^{1}, \ldots, \text{LD-DTW}_{\mathrm{I}}^{K}), \\
\boldsymbol{F\text{-}DTW}_{\mathrm{D}} &= \{\text{F-DTW}_{\mathrm{D}}^{v}\}_{v=1}^{2} \\
&= (\text{LM-DTW}_{\mathrm{D}}, \text{LD-DTW}_{\mathrm{D}}).
\end{aligned}
$$

#### *b: SCORE-LEVEL FUSION*

To achieve score-level fusion, we built two SVM classifiers using $\boldsymbol{F\text{-}DTW}_{\mathrm{I}}$ and $\boldsymbol{F\text{-}DTW}_{\mathrm{D}}$, respectively, and obtained the final score by fusing the classifier scores.

SVM [14] is a well-known machine learning classifier that is commonly used in writer and signature verification systems [4], [5], [20]. Geometrically, an SVM constructs a maximum-margin hyperplane based on the statistical learning theory principle of structural risk minimization.

When building an SVM model, we used positive instances (the intra-user variations between the target user's mean signature template set and the reference set) and negative instances (the inter-user variations between the target user's mean signature template set and the other users' mean signature template sets) for each user. Based on our preliminary results, we used a linear SVM with the $L_2$-norm penalty and the squared hinge loss with a cost-sensitive learning method to deal with imbalanced class distributions. A grid search is used to fine-tune the SVM parameter (i.e., the penalty constant $C$).

During the enrollment phase, we built two SVM classifiers using $\boldsymbol{F\text{-}DTW}_{\mathrm{I}}$ and $\boldsymbol{F\text{-}DTW}_{\mathrm{D}}$ (namely, $\text{SVM}_{\mathrm{I}}$ and $\text{SVM}_{\mathrm{D}}$, respectively). When we feed the $\boldsymbol{F\text{-}DTW}_{\mathrm{I}}$ and $\boldsymbol{F\text{-}DTW}_{\mathrm{D}}$ of a query sample into $\text{SVM}_{\mathrm{I}}$ and $\text{SVM}_{\mathrm{D}}$, we obtain the confidence scores, $\text{Score}_{\mathrm{I}}$ and $\text{Score}_{\mathrm{D}}$, which are proportional to the signed distance of that sample to the hyperplane.

Finally, we obtained a final score, $\text{Score}_{\mathrm{ID}}$, by combining the two scores as follows:

$$
\text{Score}_{\mathrm{ID}} = \text{Score}_{\mathrm{I}} + \text{Score}_{\mathrm{D}}. \tag{18}
$$

### *E. OUTPUT*

In the verification phase, the system outputs an accept or reject result based on whether the extent of dissimilarities is below or above the user-specific threshold after evaluating the scores between the purported user's mean signature template sets and test samples. We defined the threshold in this study by examining the equal error rate (EER) (Section IV-A2).

## IV. EXPERIMENTS

### *A. METHODS*

In real scenarios, skilled forgery detection is a challenging task, particularly for forensic document examiners (FDEs) [6], [7].

To overcome such challenges, we conducted experiments using the public online signature datasets: SVC2004 Task1/Task2 [16] and MCYT-100 [17]. These datasets contain various stylized signatures with highly skilled forgeries obtained from other contributors with enough training time to produce valid forgeries. Because these scenarios correspond to the addressed challenge, we used the three datasets in these experiments.

#### 1) SIGNATURE DATASETS

##### *a: SVC2004 Task1 AND Task2*

The SVC2004 Task1 and Task2 datasets contain 1,600 signatures, including Western and Asian signatures from 40 users (for a total of 3,200 signatures from 80 users). Both datasets include 20 genuine signatures and 20 skillfully forged signatures for each user. To avoid privacy concerns, the original writers were advised to provide simple, invented signatures as genuine after a sufficient amount of practice. SVC2004 Task1 includes horizontal and vertical pen coordinates, in addition to time stamps and pen up/down status, all of which are captured with a digitizing tablet at a sampling rate of 100 Hz. However, the SVC2004 Task2 includes pen pressure, azimuth, and inclination signals. Only six of the seven function-based features (Section III-C), excluding the pen pressure feature, are derived from the SVC2004 Task1.

##### *b: MCYT-100*

The MCYT-100 dataset comprises 5,000 Western signatures gathered from 100 users. The data set contains horizontal and vertical pen coordinates, pressure, azimuth, and inclination with time stamps, all of which are captured by a digitizing tablet at a sampling rate of 100 Hz. Each user is represented by 25 genuine signatures and 25 skillfully forged signatures.

#### 2) EVALUATION

Finally, we assessed the signature verification performance by examining the EER with a user-dependent threshold in which the false rejection and false acceptance rates are equal. We selected $N = 5$ genuine signatures at random as the reference set in each experiment based on previous studies and real-world scenarios. When constructing an SVM model for score-level fusion, we used 5 positive and 39 negative instances on the SVC2004 Task1 and Task2; 5 positive and 99 negative instances on the MCYT-100. For the test samples in the verification phase, the remaining 15 genuine signatures and 20 skillfully forged signatures were used on the SVC2004 Task1 and Task2; the remaining 20 genuine signatures and 25 skillfully forged signatures were used on the MCYT-100. To prevent selection bias, we repeated all experiments five times on these three datasets. Finally, we obtained the average EERs.

### *B. RESULTS*

#### 1) OVERALL PERFORMANCE

To confirm the effectiveness of the proposed method in template matching, we compared multiple combinations of templates and distance measures using three datasets under the same experimental conditions (Section IV-A).

The sets of templates and distance measures used are listed below:

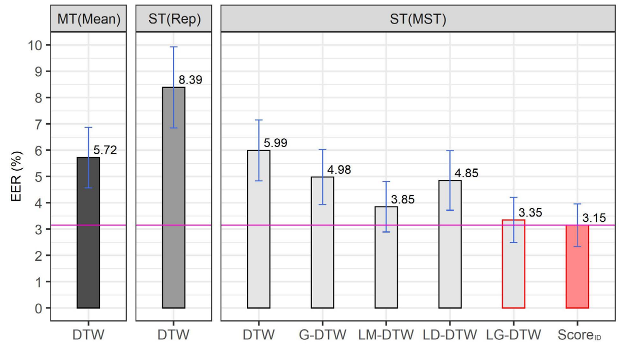
- Template strategies:
  (1) "MT(Mean)":
  a multiple-template strategy with a mean measure, which was reported to perform best among statistical measures [8], after the distances between a test sample and all references were calculated.
  (2) "ST(Rep)": a single-template strategy in which a representative template set was directly selected from the reference set using the minimum average distance measures from other samples.
  (3) "ST(MST)": a single-template strategy with a mean signature template set created using EB-DBA.
- Distance measures:
  (1) "DTW":
  the traditional DTW [23] with no weighting for the cost function.
  (2) "G-DTW":
  the previous DTW [24], applying global weighting to combine the multiple DTWs through GB.
  (3) "LM-DTW":
  the recent DTW [11] with the applied MMP-based local stability sequence as the weights for the cost function.
  (4) "LD-DTW":
  the recent DTW [13] with the applied DMP-based local stability sequence as the weights for the cost function.
  (5) "LG-DTW":
  the relevant measure [15], applying local and global weighting to the DTW.
  After obtaining the F-DTW, we calculated the global weighting factors estimate through GB: $\{\alpha_u\}_{u=1}^{2K}$ for the $\{\text{F-DTW}_{\text{I}}^{u}\}_{u=1}^{2K}$ and $\{\beta_v\}_{v=1}^{2}$ for the $\{\text{F-DTW}_{\text{D}}^{v}\}_{v=1}^{2}$, each of which satisfies $\sum_{u=1}^{2K}\alpha_u = 1$ and $\sum_{v=1}^{2}\beta_v = 1$, followed by computing the LG-DTW:

$$\text{LG-DTW} = \sum_{u=1}^{2K}\left(\alpha_u \times \text{F-DTW}_{\text{I}}^{u}\right) + \sum_{v=1}^{2}\left(\beta_v \times \text{F-DTW}_{\text{D}}^{v}\right). \quad (19)$$

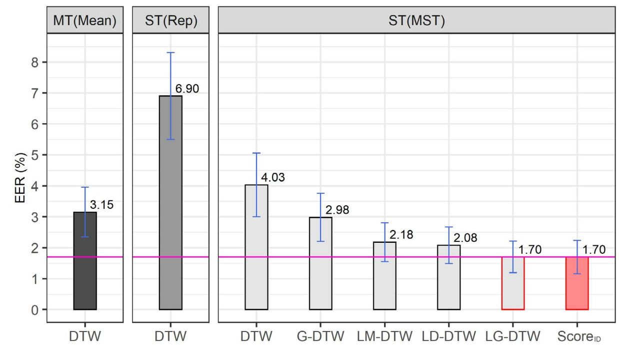
  (6) "Score$_{\text{ID}}$":
  the proposed method, applying the single-template technique to obtain a discriminative fused score through SVMs constructed using F-DTW.

Figure 4 shows the overall performance of the proposed method in terms of EER. As shown in Fig. 4, we deduced the following results:

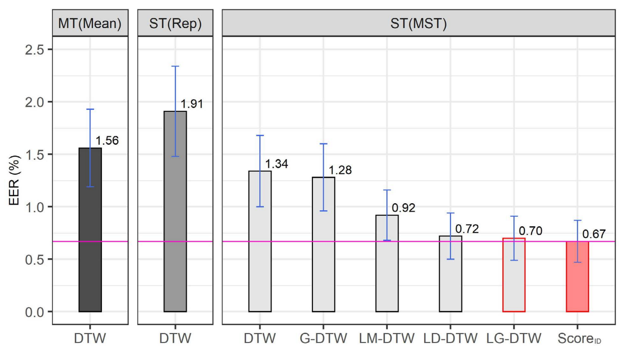
- Among DTW measures, performance using mean signature templates ("ST(MST)") is considerably better than the conventional "ST(Rep)" and competitive with the multiple-template strategy ("MT(Mean)").




(a) SVC2004 Task1 dataset.




(b) SVC2004 Task2 dataset.




(c) MCYT-100 dataset.

**FIGURE 4. Overall performance of the proposed method. Each error bar represents the standard error (SE) of the average EER.**

- In "ST(MST)," the performance is further improved by applying weighting schemes for DTW.
- Among weighting schemes for DTW, LG-DTW and Score$_{\text{ID}}$, both of which use F-DTW, outperform the independent use of G-DTW, LM-DTW, and LD-DTW. Note that the independent use of LM-DTW and LD-DTW reveals data dependency in performance among datasets; thus, it is rational to use both LM-DTW and LD-DTW in the proposed method.
- Among methods using F-DTW, performance using the score-level fusion with SVM (Score$_{\text{ID}}$) is better than the global weighting scheme with GB (LG-DTW).
- To summarize, the proposed single-template strategy (Score$_{\text{ID}}$ with "ST(MST)") achieves the lowest EERs across all datasets.

To confirm the statistical significance between the proposed $\text{Score}_{\text{ID}}$ and other seven methods, we applied the statistical hypothesis tests. We used the Matched-Pairs test [35] along with the Holm method [36] after confirming that the global hypothesis tests were significant on all datasets (i.e., the Friedman test [37] with a significant level of less than 0.001). The Matched-Pairs test determines whether the difference in errors between two methods tested on the same dataset for equivalent subjects is statistically significant. The Holm method is used to adjust the predefined significance level for multiple comparisons.

Consequently, the $\text{Score}_{\text{ID}}$ outperformed four methods (i.e., all three DTWs and G-DTW) on all datasets at a significant level of less than 0.05 (i.e., $1.58 \times 10^{-8} \leq p\text{-value} \leq 1.37 \times 10^{-3}$). As a result, we can conclude that there is a significant difference between the proposed method's results and those of other four methods.

However, we were unable to confirm the statistical significance of $\text{Score}_{\text{ID}}$ and LM-DTW or LD-DTW, which is dependent on datasets. This result shows that the individual use of LM-DTW and LD-DTW is susceptible to writing conditions; therefore, it is reasonable to use both LM-DTW and LD-DTW in the proposed method. Furthermore, we were unable to confirm the statistical significance of $\text{Score}_{\text{ID}}$ and LG-DTW, both of which use F-DTW across all datasets. This result shows that the proposed F-DTW has sufficient discriminative power in all fusion schemes; thus, it is appropriate to use score-level fusion with linear SVM in $\text{Score}_{\text{ID}}$, which requires fewer parameters and has a lower computational complexity than the global weighting scheme with GB in LG-DTW.

These results confirm that the proposed single-template strategy provides an effective template-matching approach for online signature verification.

### 2) COMPARATIVE ANALYSIS OF WEIGHTED DTW

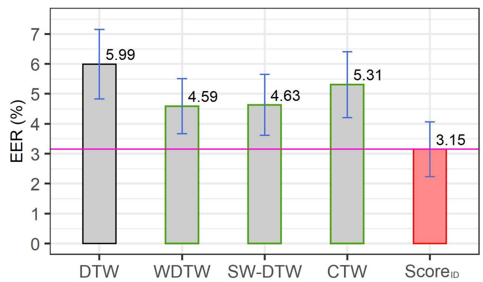
To confirm the effectiveness of the proposed method in terms of distance measures, we compared the performance with the previous weighted DTW methods [27], [28], [32] by applying the single-template strategy under the same experimental conditions.

The baselines of the weighted DTW are shown below:

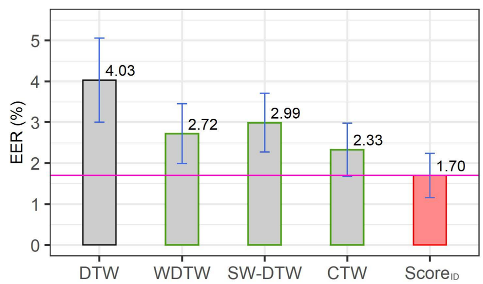
- "WDTW":
  applying the modified logistic weight function as the weight for the cost function [27].
- "SW-DTW":
  modifying the cost function by incorporating a weighted average of the neighboring distances using a window function with a width $\delta \in \mathbb{N}$ while weighting with a constant $\alpha \in [0, 1]$ between the cost in amplitude and first-order derivative [28].
- "CTW":
  combining DTW and CCA to allow feature weighting/selection and dimensionality reduction mechanisms [32].

Note that WDTW and SW-DTW were calculated using dependent warping based on the findings [8], [26] and their parameters were selected as per previous studies.

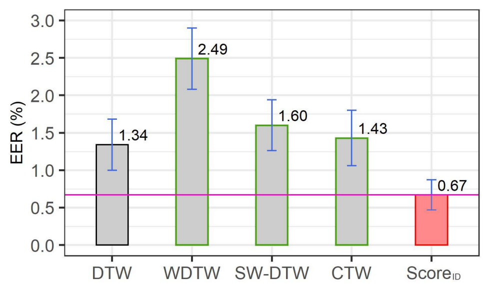
Figure 5 compares the EERs of the proposed method with the previous weighted DTW using the single-template strategy with the mean signature templates. As a baseline, we displayed the results of the conventional DTW in Fig. 5, in which we confirmed the statistically significant differences with $\text{Score}_{\text{ID}}$ (Section IV-B1). As shown in the figure, the proposed method ($\text{Score}_{\text{ID}}$) provides data independence and the lowest EERs compared with conventional methods for all datasets.




(a) SVC2004 Task1 dataset.




(b) SVC2004 Task2 dataset.




(c) MCYT-100 dataset.

**FIGURE 5. Comparison between the proposed $\text{Score}_{\text{ID}}$ and other conventional weighted DTWs using the single-template strategy with the mean signature templates. Each error bar represents the SE of the average EER.**

Following the previous experiments, we used statistical hypothesis tests to confirm the statistical significance of the performance between the proposed method and recent WDTW, SW-DTW, and CTW (Section IV-B1). Consequently, the proposed method outperformed all recent methods on all datasets at a significant level of less than 0.05, whereas most of the experiments were outperformed at a

**TABLE 1.** Comparison between the proposed method and other systems for SVC2004 Task1 dataset.

| Method | #References | EER (%) |
|---|---|---|
| Wavelet packet-based method [38] | 5 | 6.65 |
| Signatures turning angle representation using relative longest common subsequence matching [39] | 5 | 5.33 |
| **Single-template strategy (Proposed)** | **5** | **3.15** |
| Hilbert scanning patterns and Gaussian mixture models [40] | 10 | 6.08 |
| **Single-template strategy (Proposed)** | **10** | **3.14** |

**TABLE 2.** Comparison between the proposed method and other systems for the SVC2004 Task2 dataset.

| Method | #References | EER (%) |
|---|---|---|
| Dynamic time functions and HMMs [19] | 5 | 6.90 |
| SVMs with the longest common subsequences kernel function [20] | 5 | 6.84 |
| Enhanced contextual DTW-based system using vector quantization [41] | 5 | 2.73 |
| SynSig2Vec using a lightweight 1D CNN trained with synthesized signatures [42] | 5 | 2.63 |
| Discriminative feature selection and DTW with signature curve constraint [43] | 5 | 2.60 |
| DTW and warping path-based features [44] | 5 | 2.53 |
| Two-stage method using shape contexts and function features [45] | 5 | 2.39 |
| SynSig2Vec using a novel 1D CNN trained with a learning-by-synthesis method [21] | 5 | 2.08 |
| **Single-template strategy (Proposed)** | **5** | **1.70** |
| Semiparametric method based on discrete cosine transform and sparse representation [46] | 10 | 3.98 |
| Template selection and DTW [47] | 10 | 2.84 |
| DTW and warping path-based features [44] | 10 | 2.79 |
| **Single-template strategy (Proposed)** | **10** | **1.33** |

**TABLE 3.** Comparison between the proposed method and other systems for the MCYT-100 dataset.

| Method | #References | EER (%) |
|---|---|---|
| Signature partitioning and the weights of importance for selected partitions [48] | 5 | 4.88 |
| Histogram-based features and Manhattan distance [49] | 5 | 4.02 |
| Combination of global and regional features [50] | 5 | 3.69 |
| DTW and sigma-lognormal analysis [51] | 5 | 3.56 |
| Information divergence-based matching strategy [18] | 5 | 3.16 |
| SM-DTW using distance normalization [29] | 5 | 3.09 |
| Interval-valued symbolic representation with writer dependent parameters [52] | 5 | 2.2 |
| Discriminative feature selection and DTW with signature curve constraint [43] | 5 | 2.17 |
| RNNs for representation learning in the DTW framework [53] | 5 | 1.81 |
| Enhanced contextual DTW-based system using vector quantization [41] | 5 | 1.55 |
| DTW and warping path-based features [44] | 5 | 1.15 |
| SynSig2Vec using a lightweight 1D CNN trained with synthesized signatures [42] | 5 | 0.93 |
| **Single-template strategy (Proposed)** | **5** | **0.67** |

significance level of less than 0.001 (i.e., $2.60 \times 10^{-12} \leq p$-value $\leq 1.99 \times 10^{-2}$). Therefore, we can conclude that there is a significant difference between the results of the proposed method and each of the recent WDTW, SW-DTW, and CTW.

These results confirm that the proposed method provides an effective measure, particularly for the single-template strategy in online signature verification.

### 3) COMPARATIVE ANALYSIS OF STATE-OF-THE-ART SYSTEMS

To assess the effectiveness of the proposed single-template strategy in online signature verification, we compared the results of the proposed method's EERs with those of state-of-the-art systems.

Tables 1–3 present the results obtained using the SVC2004 Task1/Task2 and MCYT-100, respectively, where only genuine signatures for the enrollment phase and genuine and skillfully forged signatures for the verification phase.

As per the previous experimental results, we displayed the EERs of the proposed method as a representative of the single-template strategy in these tables (Section IV-B1). Note that the comparative analysis of the SVC2004 Task1/Task2 datasets was set up as experiments using not only $N = 5$ but also $N = 10$ as the reference signatures for fair comparisons with the previous studies (Tables 1 and 2).

In all datasets, the proposed single-template strategy outperforms other recent literature systems as per these tables. Even while investigating skilled forgery scenarios, the results confirm the effectiveness of the proposed method for online signature verification.

## V. DISCUSSION

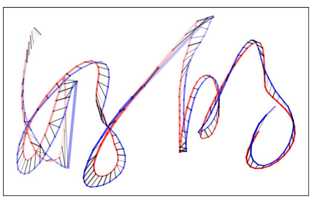
In this study, we proposed a novel single-template strategy that, when compared to the multiple-template strategy, provides lower calculation complexity while providing higher verification performance. The primary contributions of this study are summarized as follows:

- We adopted the mean signature template creation method with EB-DBA to incorporate intra-user variations within the reference samples.
- It provides locally weighted DTWs using both MMPs and DMPs (i.e., LM-DTW and LD-DTW, respectively) derived from the intra-user variations between the mean signature template and reference samples for independent and dependent warping to incorporate detailed and flexible local stability information and to effectively minimize intra-class discrepancies.
- To improve inter-user variability, it employs multiple fusion strategies: representation-level fusion, which concatenates LM-DTW and LD-DTW into a single vector, F-DTW, for each warping; and score-level fusion, which combines each score via SVMs constructed using F-DTW for each warping.

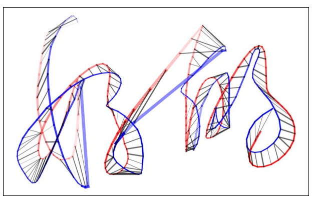
Unlike recent black-box modeling strategies such as deep learning algorithms that require high computational complexity and many training samples, the proposed approach is superior, particularly in forensic situations with limited available data [4]–[7].

The proposed method depends on explainable stepwise methods that support FDEs to explore their differences and similarities and explain the rationale behind their assessment of legal professionals in the decision-making process. For example, it can provide detailed matching between the mean signature template and a query signature (Figs. 1 and 2).

Furthermore, the proposed method can provide visual investigation tools to assist FDEs in forensic analysis in an explainable manner. For example, we can see F-DTW$_\mathrm{D}$ matching between an EB-DBA mean signature template and query signatures in the spatiotemporal domain (Fig. 6). In this diagram, the greater the distance with a higher local weighting between the matching points, the greater the likelihood that the query signature is a forgery.



(a) A mean signature template and a genuine signature.



(b) A mean signature template and a skilled forgery

**FIGURE 6. Examples of F-DTW$_\mathrm{D}$ matching (black lines) between a mean signature template obtained through EB-DBA (red strokes) and genuine or forged signatures (blue strokes) in the spatiotemporal domain. The alignment line between the matching points becomes thicker according to the fused local weighting estimated with $LM_\mathrm{D}$ and $LD_\mathrm{D}$.**

Consequently, it is particularly useful for applications, such as forensics and security, in which fairness, accountability, and transparency are important.

## VI. CONCLUSION

To increase the performance of online signature verification, we devised a unique single-template technique for DTW based on mean signature template sets and fusion strategies of multiple local weighting and warping schemes.

During the enrollment phase, we used EB-DBA to generate user-specific mean signature template sets while accounting for intra-user heterogeneity across reference samples. Then, for independent and dependent DTW, we computed local weighting estimates by analyzing MMPs and DMPs between mean signature template sets and reference samples, respectively, to incorporate detailed and flexible local stability information and effectively minimize intra-class discrepancies. To improve inter-user variability, we used the representation-level fusion to concatenate LM-DTW and LD-DTW calculated with MMPs and DMPs in a single vector, F-DTW, for each warping, followed by the score-level fusion to combine each score through SVMs constructed using F-DTW for each warping.

The experimental results on the public online signature datasets, SVC2004 Task1/Task2 and MCYT-100, proved the usefulness of the suggested method for online signature verification.

The proposed explainable stepwise strategy can bridge and compensate for the biometrics–forensics divide.

## REFERENCES


[1] R. Plamondon, G. Pirlo, E. Anquetil, C. Rémi, H.-L. Teulings, and M. Nakagawa, "Personal digital bodyguards for e-security, e-learning and e-health: A prospective survey," *Pattern Recognit.*, vol. 81, pp. 633–659, Sep. 2018.

[2] M. Diaz, M. A. Ferrer, D. Impedovo, M. I. Malik, G. Pirlo, and R. Plamondon, "A perspective analysis of handwritten signature technology," *ACM Comput. Surv.*, vol. 51, no. 6, pp. 117:1–117:39, 2019.

[3] D. Impedovo and G. Pirlo, "Automatic signature verification in the mobile cloud scenario: Survey and way ahead," *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 1, pp. 554–568, Jan./Mar. 2021.

[4] M. Okawa and K. Yoshida, "Offline writer verification based on forensic expertise: Analyzing multiple characters by combining the shape and advanced pen pressure information," *Jpn. J. Forensic Sci. Technol.*, vol. 22, no. 2, pp. 61–75, 2017.

[5] M. Okawa, "Synergy of foreground–background images for feature extraction: Offline signature verification using Fisher vector with fused KAZE features," *Pattern Recognit.*, vol. 79, pp. 480–489, Jul. 2018.

[6] C. M. Deviterne-Lapeyre, "Interpol review of questioned documents 2016–2019," *Forensic Sci. Int., Synergy*, vol. 2, pp. 429–441, Jan. 2020.

[7] D. Mazzolini, P. Mignone, P. Pavan, and G. Vessio, "An easy-to-explain decision support framework for forensic analysis of dynamic signatures," *Forensic Sci. Int., Digit. Invest.*, vol. 38, Sep. 2021, Art. no. 301216.

[8] M. Okawa, "Template matching using time-series averaging and DTW with dependent warping for online signature verification," *IEEE Access*, vol. 7, pp. 81010–81019, 2019.

[9] M. Okawa, "Online signature verification using single-template matching through locally and globally weighted dynamic time warping," *IEICE Trans. Inf. Syst.*, vol. E103.D, no. 12, pp. 2701–2708, 2020.

[10] M. Okawa, "Modified dynamic time warping with local and global weighting for online signature verification," in *Proc. IEEE 3rd Global Conf. Life Sci. Technol. (LifeTech)*, Mar. 2021, pp. 124–125.

[11] M. Okawa, "Online signature verification using a single-template strategy with mean templates and local stability-weighted dynamic time warping," in *Proc. IEEE 11th Int. Workshop Comput. Intell. Appl. (IWCIA)*, Nov. 2019, pp. 83–88.

[12] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.

[13] M. Okawa, "Time-series averaging and local stability-weighted dynamic time warping for online signature verification," *Pattern Recognit.*, vol. 112, Apr. 2021, Art. no. 107699.

[14] V. N. Vapnik, *Statistical Learning Theory*. Hoboken, NJ, USA: Wiley, 1998.

[15] M. Okawa, "Multiple local weighting scheme for dynamic time warping in online signature verification," in *Proc. IEEE 4th Global Conf. Life Sci. Technol. (LifeTech)*, Mar. 2022, pp. 280–281.
[16] D.-Y. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, and G. Rigoll, "SVC2004: First international signature verification competition," in *Proc. Int. Conf. Biometric Authentication (ICBA)*, in Lecture Notes in Computer Science, vol. 3072, 2004, pp. 16–22.
[17] 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, and Q.-I. Moro, "MCYT baseline corpus: A bimodal biometric database," *IEE Proc., Vis., Image Signal Process.*, vol. 150, no. 6, pp. 395–401, 2003.
[18] L. Tang, W. Kang, and Y. Fang, "Information divergence-based matching strategy for online signature verification," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 4, pp. 861–873, Apr. 2018.
[19] J. Fierrez, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez, "HMM-based on-line signature verification: Feature extraction and signature modeling," *Pattern Recognit. Lett.*, vol. 28, no. 16, pp. 2325–2334, 2007.
[20] C. Gruber, T. Gruber, S. Krinninger, and B. Sick, "Online signature verification with support vector machines based on LCSS kernel functions," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 4, pp. 1088–1100, Aug. 2010.
[21] S. Lai, L. Jin, Y. Zhu, Z. Li, and L. Lin, "SynSig2Vec: Forgery-free learning of dynamic signature representations by sigma lognormal-based synthesis," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Jun. 8, 2021, doi: 10.1109/TPAMI.2021.3087619.
[22] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Exploring recurrent neural networks for on-line handwritten signature biometrics," *IEEE Access*, vol. 6, pp. 5128–5138, 2018.
[23] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-26, no. 1, pp. 43–49, Feb. 1978.
[24] M. Okawa, "Online signature verification using single-template matching with time-series averaging and gradient boosting," *Pattern Recognit.*, vol. 102, Jun. 2020, Art. no. 107227.
[25] A. Mueen and E. Keogh, "Extracting optimal performance from dynamic time warping," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 2129–2130.
[26] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. Keogh, "Generalizing DTW to the multi-dimensional case requires an adaptive approach," *Data Mining Knowl. Discovery*, vol. 31, no. 1, pp. 1–31, 2017.
[27] Y.-S. Jeong, M. K. Jeong, and O. A. Omitaomu, "Weighted dynamic time warping for time series classification," *Pattern Recognit.*, vol. 44, no. 9, pp. 2231–2240, Sep. 2011.
[28] D. Folgado, M. Barandas, R. Matias, R. Martins, M. Carvalho, and H. Gamboa, "Time alignment measurement for time series," *Pattern Recognit.*, vol. 81, pp. 268–279, Sep. 2018.
[29] A. Parziale, M. Diaz, M. A. Ferrer, and A. Marcelli, "SM-DTW: Stability modulated dynamic time warping for signature verification," *Pattern Recognit. Lett.*, vol. 121, pp. 113–122, Apr. 2019.
[30] D. Impedovo, G. Pirlo, M. Diaz, and M. A. Ferrer, "Weighted direct matching points for user stability model in multiple domains: A proposal for on-line signature verification," in *Proc. 15th Int. Conf. Document Anal. Recognit. (ICDAR)*, Sep. 2019, pp. 1320–1325.
[31] D. Impedovo, G. Pirlo, and L. Sarcinella, "Signatures' stability evaluation in a multi-device scenario," in *Proc. 17th Int. Conf. Frontiers Handwriting Recognit. (ICFHR)*, Sep. 2020, pp. 367–372.
[32] F. Zhou and F. De la Torre, "Generalized canonical time warping," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 279–294, Feb. 2016.
[33] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey, A. Ragni, V. Valtchev, P. Woodland, and C. Zhang, *The HTK Book (for HTK Version 3.5, Documentation Alpha Version)*. Cambridge, U.K.: Cambridge Univ. Press, 2015.
[34] F. Petitjean, A. Ketterlin, and P. Gançarski, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern Recognit.*, vol. 44, no. 3, pp. 678–693, Mar. 2011.
[35] L. Gillick and S. J. Cox, "Some statistical issues in the comparison of speech recognition algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 1989, pp. 532–535.
[36] S. Holm, "A simple sequentially rejective multiple test procedure," *Scand. J. Statist.*, vol. 6, no. 2, pp. 65–70, 1979.
[37] M. Friedman, "A comparison of alternative tests of significance for the problem of *m* rankings," *Ann. Math. Statist.*, vol. 11, no. 1, pp. 86–92, 1940.
[38] K. Wang, Y. Wang, and Z. Zhang, "On-line signature verification using wavelet packet," in *Proc. Int. Joint Conf. Biometrics (IJCB)*, Oct. 2011, pp. 1–6.
[39] K. Barkoula, G. Economou, and S. Fotopoulos, "Online signature verification based on signatures turning angle representation using longest common subsequence matching," *Int. J. Document Anal. Recognit.*, vol. 16, no. 3, pp. 261–272, 2013.
[40] A. Ahrary, H.-J. Chiang, and S.-I. Kamata, "On-line signature matching based on Hilbert scanning patterns," in *Proc. Int. Conf. Biometrics*, 2009, pp. 1190–1199.
[41] A. Sharma and S. Sundaram, "An enhanced contextual DTW based system for online signature verification using vector quantization," *Pattern Recognit. Lett.*, vol. 84, pp. 22–28, Dec. 2016.
[42] S. Lai, L. Jin, L. Lin, Y. Zhu, and H. Mao, "SynSig2Vec: Learning representations from synthetic dynamic signatures for real-world verification," in *Proc. 34th AAAI Conf. Artif. Intell. (AAAI)*, 2020, pp. 735–742.
[43] X. Xia, X. Song, F. Luan, J. Zheng, Z. Chen, and X. Ma, "Discriminative feature selection for on-line signature verification," *Pattern Recognit.*, vol. 74, pp. 422–433, Feb. 2018.
[44] A. Sharma and S. Sundaram, "On the exploration of information from the DTW cost matrix for online signature verification," *IEEE Trans. Cybern.*, vol. 48, no. 2, pp. 611–624, Feb. 2018.
[45] Y. Jia, L. Huang, and H. Chen, "A two-stage method for online signature verification using shape contexts and function features," *Sensors*, vol. 19, no. 8, Apr. 2019, Art. no. 1808.
[46] Y. Liu, Z. Yang, and L. Yang, "Online signature verification based on DCT and sparse representation," *IEEE Trans. Cybern.*, vol. 45, no. 11, pp. 2498–2511, Nov. 2014.
[47] N. Liu and Y. Wang, "Template selection for on-line signature verification," in *Proc. 19th Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2008, pp. 1–4.
[48] K. Cpałka, M. Zalasiński, and L. Rutkowski, "A new algorithm for identity verification based on the analysis of a handwritten dynamic signature," *Appl. Soft Comput.*, vol. 43, pp. 47–56, Jun. 2016.
[49] N. Sae-Bae and N. Memon, "Online signature verification on mobile devices," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 6, pp. 933–947, Jun. 2014.
[50] Z. Xia, T. Shi, N. N. Xiong, X. Sun, and B. Jeon, "A privacy-preserving handwritten signature verification method using combinational features and secure kNN," *IEEE Access*, vol. 6, pp. 46695–46705, 2018.
[51] A. Fischer and R. Plamondon, "Signature verification based on the kinematic theory of rapid human movements," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 2, pp. 169–180, Apr. 2017.
[52] D. S. Guru, K. Manjunatha, S. Manjunath, and M. Somashekara, "Interval valued symbolic representation of writer dependent features for online signature verification," *Expert Syst. Appl.*, vol. 80, pp. 232–243, Sep. 2017.
[53] S. Lai and L. Jin, "Recurrent adaptation networks for online signature verification," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 6, pp. 1624–1637, Jun. 2019.

**MANABU OKAWA** (Member, IEEE) received the B.S. degree in science from Tokyo Metropolitan University, Tokyo, Japan, in 2001, the M.S. degree in engineering from Shinshu University, Nagano, Japan, in 2008, and the Ph.D. degree in engineering from the University of Tsukuba, Tsukuba, Japan, in 2014. From 2016 to 2019, he was a Part-Time Lecturer at the University of Tsukuba. He is with the Forensic Document Section of Forensic Science Laboratory, Metropolitan Police Department, Tokyo. His research interests include biometrics, forensics, and data science.