Tracking Differentiator-based Multiview Dilated Characteristics for Time Series Classification

1st Changchun He
Control and Simulation Center
Harbin Institute of Technology
Harbin, China
22b904060@stu.hit.edu.cn

2nd Xin Huo*
Control and Simulation Center
Harbin Institute of Technology
Harbin, China
huoxin@hit.edu.cn

***Abstract*—In industrial production, obtaining the discrete status of equipment from massive time series data has become an urgent demand, which places higher requirements on the time series classification (TSC) algorithm. The feature-based TSC method achieves interpretable and potential classification capability by extracting meaningful descriptive statistical features. Limitations exist in current feature-based TSC algorithms, hindering the acquisition of diverse statistical features due to a lack of knowledge and insufficient research on processing redundant features, thereby limiting performance enhancements. Therefore, in this article, a novel feature-based TSC algorithm tracking differentiator-based multiview dilated characteristics (TD-MVDC) is proposed. The innovative introduction of a tracking differentiator combined with dilation mapping as a preprocessor into the feature-based TSC method is proposed to improve feature diversity efficiently. Ensemble feature selection based on filter feature selectors with different store ratios is designed to generate multiview features to enhance feature stability quickly. Linear classifiers and hard voting to fastly classify and integrate multiview features to increase classification performance robustly. Finally, comparative and ablative experiments are conducted between TD-MVDC and the representative feature-based TSC algorithms on the extensively compared UCR archive to verify the effectiveness of the proposed algorithm.**

***Index Terms*—time series classification, tracking differentiator, dilation mapping, statistical features, ensemble feature selection**

I. Introduction

To monitor the status of industrial production, sensor systems are installed on equipment to collect and store massive time series (TS) signals, which serve as the data basis for TS processing [1]–[3]. The widespread attention received by time series classification (TSC) methods is attributed to the urgent industrial demand for improving equipment reliability via TS analysis [4]–[7]. The feature-based TSC methods have become a hot topic in recent TSC research because of their simple architecture, ease of parallelization, and potential ability [8].

The feature-based TSC methods adopt the series-to-vector structure, generally extracting a vast number of statistical features from the whole TS and then inputting the feature vector into the classifier to predict the label [9]. The representative algorithms, namely Time Series Feature Extraction based on Scalable Hypothesis Tests (TSFresh) [10], 22 Canonical Time Series Characteristics (Catch22) [11], Generalized Signatures [12], and Fresh Pipeline with Rotation Forest Classifier (FreshPRINCE) [9], are differentiated based on the features extracted and the classifier utilized. The features with clear and descriptive meanings are extracted via the feature-based TSC algorithms, thus improving interpretability and acceptance by users in the non-machine learning community.

Shortcomings are observed in existing feature-based TSC algorithms, such as a single approach to increase feature diversity. At a certain level of feature scale, extracting new features becomes challenging due to expertise and statistical knowledge limitations. The series mapping is a flexible method to improve feature diversity, among which differential transform is widely used, and dilation mapping is a rising star proposed by Schaefer and Leser [13]. Differential transform is contaminated by noise easily, and the tracking differentiator (TD) proposed by Han suppresses the adverse effects of noise adjustably [14]. Introducing the dilation mechanism into the feature-based TSC method is natural and interesting [8], as it is the critical component of state-of-the-art TSC algorithms for identifying patterns from multiple scales [13], [15].

Another limitation is that there are redundant and irrelevant features in the massive features extracted. It is necessary to introduce the feature selection method, which removes redundant features and indicates the importance of features. Ensemble feature selection is considered a promising feature selection method and integrates filter, wrapper, and embedded traditional feature selection algorithms to improve feature stability and robustness for parameters [16], [17].

Therefore, the motivation of this article is to introduce series mapping and ensemble feature selection to enhance feature diversity and stability, thereby efficiently improving the classification performance of the feature-based TSC algorithm. A new tracking differentiator-based multiview dilated characteristics (TD-MVDC) is proposed for TSC, including three processes: various feature extraction, ensemble feature selection, and multiview ensemble classification.

The contributions of the article are as follows:

1) A novel feature-based TSC algorithm, TD-MVDC, is proposed, in which the series mapping and ensemble feature selection are introduced into the feature-based TSC method, improving the classification performance.

This work was supported in part by the National Natural Science Foundation of China under Grant 62373128, and in part by the Natural Science Foundation of Heilongjiang Province of China under Grant LH2021F025.

2) A new differential dilation series mapping algorithm is presented as the series preprocessor, which combines the TD with the ability to suppress noise and dilation mapping with the capability to recognize patterns from multiple scales, increasing the feature diversity effectively.
3) An ensemble feature selection algorithm is designed to integrate filter feature selection methods using different store ratios, enhancing feature stability quickly.

The rest of article is organized as follows: The related feature-based TSC and series mapping algorithms are introduced in Section II. In Section III, a novel feature-based TSC algorithm TD-MVDC is proposed, and in Section IV, the effectiveness of TD-MVDC is verified by comparative and ablative experiments on the public UCR archive. The conclusion and the future work are presented in Section V.

II. Related Work

A. Feature-based TSC Algorithms

The general classification block diagram of the feature-based TSC method is shown in Fig. 1, comprising two steps: feature extraction and classification. The main difference between each feature-based TSC algorithm is the type of features extracted and the classifier utilized, among which representative algorithms are introduced.

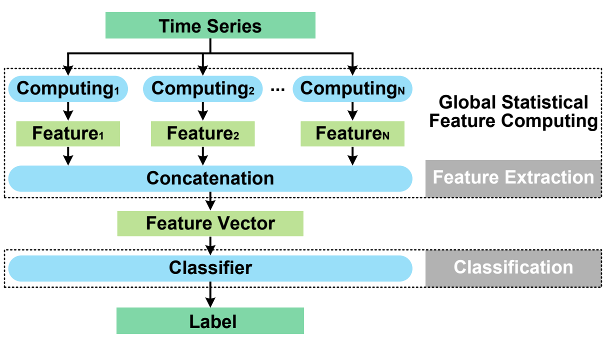

Fig. 1. Block diagram of feature-based TSC algorithms.

1) TSFresh: TSFresh is a Python package proposed by Christ et al. that extracts less than 800 TS features and has the feature selection function [10]. TSFresh uses multiple hypothesis tests to remove redundant features and then inputs the selected features into the random forest classifier [8], [10].

2) Catch22: Lubba et al. selected the most discriminative 22 canonical features of TS via pruning, evaluation, and clustering of more than 7700 TS features extracted from the highly comparative time-series analysis (hctsa) toolbox [18] and combined the decision tree classifier [11].

3) Generalised Signatures: Generalised signatures was proposed by Morrill et al. based on the rough path theory, using basepoint and time enhancements for the TS and then performing signature transform to extract features and using the random forest classifier [12].

4) FreshPRINCE: The FreshPRINCE algorithm, proposed by Middlehurst and Bagnall, differs from the original TSFresh classification algorithm in that it utilizes a pipeline structure to directly input all TSFresh features into the rotation forest classifier without performing feature selection [9].

By summarizing these representative feature-based TSC algorithms, interpretive statistical features are summarized, providing a concise and feasible TSC paradigm. However, limited by knowledge, it is difficult to extract new original statistical features and increase feature diversity directly.

B. Time Series Mapping

Mapping original TS data to other series spaces via a fast preprocessor is an effective technique, with differential and dilation mapping having promising potential.

1) Tracking Differentiator: Differential series with clear meaning contains the changing information and dynamic patterns of the original series. According to (1), the differential transform is used for $\boldsymbol{s} = (s_1, s_2, \dots, s_m)$ by many state-of-the-art TSC algorithms [13], [15], but the differential series $\boldsymbol{s^F}$ obtained by (1) is easily contaminated by noise.

$$\boldsymbol{s^F} = Difference(\boldsymbol{s}) = (s_1^F, s_2^F, \dots, s_{m-1}^F) \tag{1}$$

where $s_i^F = s_{i+1} - s_i, i = 1, 2, \dots, m-1$.

The pseudocode of TD is shown in **Algorithm 1**, where r is the velocity factor, h is the step size, and k is the filter factor. As the value of k increases, the filtering ability is strengthened, and the noise content in the differential series is reduced.

Algorithm 1 Tracking Differentiator [14]

Parameters: r (default 100), h (default 1), k (default 2)
Input: Time series $\boldsymbol{s} = (s_1, s_2, \dots, s_m)$
Output: Differential series $\boldsymbol{s^F} = (s_1^F, s_2^F, \dots, s_{m-1}^F)$
$h_0 = k \cdot h$
$\boldsymbol{x_1}[1],\ \boldsymbol{x_2}[1] = \boldsymbol{s}[1], (\boldsymbol{s}[2] - \boldsymbol{s}[1])/h$
for $i = 1, 2, \dots, m$ **do**
  $fh_i, y_i = \text{fhan}(\boldsymbol{x_1}[i] - \boldsymbol{s}[i],\ \boldsymbol{x_2}[i],\ r,\ h_0)$
  $\boldsymbol{x_1}[i+1] = \boldsymbol{x_1}[i] + h \cdot \boldsymbol{x_2}[i]$
  $\boldsymbol{x_2}[i+1] = \boldsymbol{x_2}[i] + h \cdot fh_i$
$\boldsymbol{s^F} = (y_2, y_3, \dots, y_m)/h_0$
Subfunction: $\text{fhan}(x_1, x_2, r, h_0)$
$d = r \cdot h_0$
$d_0 = h_0 \cdot d$
$y = x_1 + h_0 \cdot x_2$
$a_0 = \sqrt{d^2 + 8r|y|}$
if $|y| > d_0$: $a = x_2 + (a_0 - d) \cdot \text{sign}(y)/2$
else: $a = x_2 + y/h_0$
if $|a| > d$: $fh = -r \cdot \text{sign}(a)$
else : $fh = -r \cdot a/d$
return fh, y

2) Dilatation Mapping: The dilatation mechanism enables the feature extractor to obtain pattern information on multiple scales, increasing the receptive field. Dilatation mapping is a series mapping algorithm proposed by Schaefer and Leser that flexibly combines the dilatation mechanism without changing the existing feature extraction program [13]. The pseudocode of the dilation mapping is shown in **Algorithm 2**, where d is the dilation rate. Dilation mapping is suitable for parallel operations, using different d values to obtain dilated sequences of multiple scales.

Algorithm 2 Dilation Mapping [13]

Parameters: Dilation rate d
Input: Time series $\boldsymbol{s} = (s_1, s_2, \ldots, s_m)$
Output: Dilated time series $\boldsymbol{s^d} = (s_1^d, s_2^d, \ldots, s_m^d)$
$\boldsymbol{s^d} = [\,]$
for $k = 1, 2, \ldots, d$ **do**
  $\boldsymbol{s^d} = \text{Concatenate}([\boldsymbol{s^d}, \boldsymbol{s}[k :: d]])$
return $\boldsymbol{s^d} = (s_1^d, s_2^d, \ldots, s_m^d)$

III. Proposed Method

In this article, univariate TS with equal length, equal intervals, and real values are taken as the research object. The goal of TSC is to establish a mapping/classifier from TS $\boldsymbol{s} = (s_1, s_2, \ldots, s_m)$ to discrete label $label \in \{l_1, l_2, \ldots, l_C\}$, where m is the length of TS, and C is the number of discrete label types. The classifier is trained on the train set $TrainSet = (\boldsymbol{s_1}, \boldsymbol{s_2}, \ldots, \boldsymbol{s_n})$and its labels $\boldsymbol{L} = (label_1, label_2, \ldots, label_n)$ using supervised learning to predict the label of the unlabeled TS samples, where n is the number of train set samples.

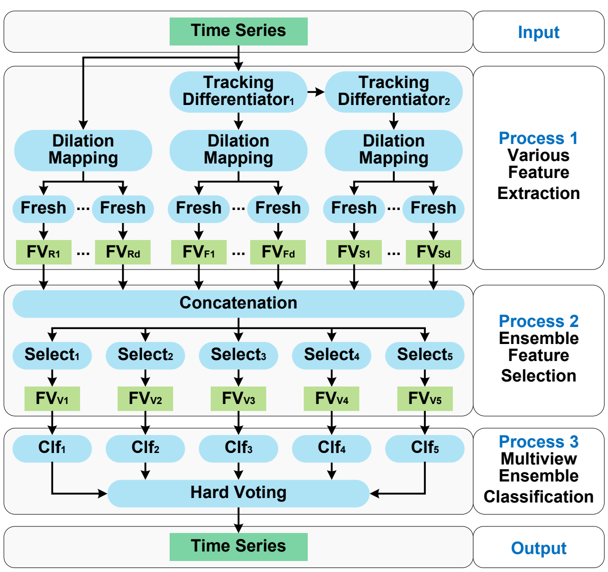

Fig. 2. The framework diagram of TD-MVDC.

A feature-based TSC algorithm TD-MVDC is proposed, which includes three processes: various feature extraction, ensemble feature selection, and multiview ensemble classification, as shown in Fig. 2. Unlike existing feature-based TSC algorithms, TD and dilation mapping are novelly introduced to improve feature diversity, and filter-based ensemble feature selection is designed to enhance feature stability. The pseudocode of TD-MVDC is shown in **Algorithm 3**, where k_1 and k_2 are the filter factors of the two TDs, d_{max} is the maximum dilation rate of dilation mapping, and $ratioList$ is the store ratio list for filter feature selectors. The training and prediction processes of TD-MVDC are shown in lines 1-16 and 17-29 of **Algorithm 3**, respectively.

Algorithm 3 Tracking Differentiator-based Multiview Dilated Characteristics (TD-MVDC)

Parameters: First-order TD filter factor k_1 (default 2),
 Second-order TD filter factor k_2 (default 2),
 Maximum dilation rate d_{max} (default 32),
 Store ratio list of filter feature selectors $ratioList$ (default $[r_1, r_2, \ldots, r_y] = [0.1, 0.2, 0.3, 0.4, 0.5]$),
Input: Train series set and its labels $\boldsymbol{S^L} = (\boldsymbol{s_1^L}, \ldots, \boldsymbol{s_a^L})$, $\boldsymbol{L}$,
 Unlabeled series set $\boldsymbol{S^U} = (\boldsymbol{s_1^U}, \ldots, \boldsymbol{s_b^U})$,
Output: Labels of unlabeled series set $\boldsymbol{L^U}$
Training:
1: $\boldsymbol{S_F^L} = \text{TD}(\boldsymbol{S^L}, k_1)$
2: $\boldsymbol{S_S^L} = \text{TD}(\boldsymbol{S_F^L}, k_2)$
3: $d_{ceil} = \text{Min}([2^{\text{int}(\log_2(m-1))-3}, d_{max}])$
4: $dList = [2^0, 2^1, \ldots, d_{ceil}]$
5: $\boldsymbol{S_{RD0}^L}, \ldots, \boldsymbol{S_{RDd}^L} = \text{Dilation}(\boldsymbol{S^L}, dList)$
6: $\boldsymbol{S_{FD0}^L}, \ldots, \boldsymbol{S_{FDd}^L} = \text{Dilation}(\boldsymbol{S_F^L}, dList)$
7: $\boldsymbol{S_{SD0}^L}, \ldots, \boldsymbol{S_{SDd}^L} = \text{Dilation}(\boldsymbol{S_S^L}, dList)$
8: **for** $k = 0, 1, \ldots, d$
9: $\boldsymbol{FV_{RDk}^L} = \text{TSFresh}(\boldsymbol{S_{RDk}^L})$
10: $\boldsymbol{FV_{FDk}^L} = \text{TSFresh}(\boldsymbol{S_{FDk}^L})$
11: $\boldsymbol{FV_{SDk}^L} = \text{TSFresh}(\boldsymbol{S_{SDk}^L})$
12: $\boldsymbol{FV^L} = \text{Concatenate}(\boldsymbol{FV_{RD0}^L}, \ldots, \boldsymbol{FV_{RDd}^L}, \boldsymbol{FV_{RF0}^L}, \ldots, \boldsymbol{FV_{RFd}^L}, \boldsymbol{FV_{RS0}^L}, \ldots, \boldsymbol{FV_{RSd}^L})$
13: $\boldsymbol{Scores} = \text{Correlation}(\boldsymbol{FV^L}, \boldsymbol{L})$
14: **for** $k = 1, 2, \ldots, y$
15: $index_k$ is the index of the top $ratio_k$ features
16: $\text{clf}_\text{k} = \text{Classifier.training}(\boldsymbol{FV^L}[:, index_k], \boldsymbol{L})$
Prediction:
17: $\boldsymbol{S_F^U} = \text{TD}(\boldsymbol{S^U}, k_1)$
18: $\boldsymbol{S_S^U} = \text{TD}(\boldsymbol{S_F^U}, k_2)$
19: $\boldsymbol{S_{RD0}^U}, \ldots, \boldsymbol{S_{RDd}^U} = \text{Dilation}(\boldsymbol{S_R^U}, dList)$
20: $\boldsymbol{S_{FD0}^U}, \ldots, \boldsymbol{S_{FDd}^U} = \text{Dilation}(\boldsymbol{S_F^U}, dList)$
21: $\boldsymbol{S_{SD0}^U}, \ldots, \boldsymbol{S_{SDd}^U} = \text{Dilation}(\boldsymbol{S_S^U}, dList)$
22: **for** $k = 0, 1, \ldots, d$
23: $\boldsymbol{FV_{RDk}^U} = \text{TSFresh}(\boldsymbol{S_{RDk}^U})$
24: $\boldsymbol{FV_{FDk}^U} = \text{TSFresh}(\boldsymbol{S_{FDk}^U})$
25: $\boldsymbol{FV_{SDk}^U} = \text{TSFresh}(\boldsymbol{S_{SDk}^U})$
26: $\boldsymbol{FV^U} = \text{Concatenate}(\boldsymbol{FV_{RD0}^U}, \ldots, \boldsymbol{FV_{RDd}^U}, \boldsymbol{FV_{RF0}^U}, \ldots, \boldsymbol{FV_{RFd}^U}, \boldsymbol{FV_{RS0}^U}, \ldots, \boldsymbol{FV_{RSd}^U})$
27: **for** $k = 1, 2, \ldots, y$
28: $\boldsymbol{L_k^U} = \text{clf}_\text{k}.\text{predict}(\boldsymbol{FV^U}[:, index_k])$
29: $\boldsymbol{L^U} = \text{HardVoting}([\boldsymbol{L_1^U}, \boldsymbol{L_2^U}, \ldots, \boldsymbol{L_y^U}])$
return $\boldsymbol{L^U}$

A. Various Feature Extraction

The motivation behind various feature extraction of TD-MVDC is to efficiently enhance feature diversity by utilizing the proposed unsupervised differential dilation mapping as a series preprocessor. Velocity and acceleration information is obtained by first-order and second-order differential mappings using TDs with adjustable filter factors. Then, the original and differential sequences are subjected to dilation mapping to introduce the dilation mechanism for recognizing multiple scale patterns. Finally, the TSFresh features extracted from each mapped series are fully concatenated into a feature vector with diversity.

In the training process, the first-order and second-order differential series are obtained using TDs with filter factors k_1 and k_2 in lines 1-2. In lines 3-4, the dilation rate list is calculated, as referenced from [15], where m is the length of TS and the dilation rate is 1 represents the original series. The original and differential sequences are dilated at multiple rates according to $dList$, in lines 5-7. In lines 8-11, less than 800 TSFresh features are extracted for each dilated series, where it is worth noting that the series are normalized before extraction. All extracted features are spliced into a feature vector in line 12. Each TS is converted into a vector containing $3 \cdot \mathrm{Len}(dList) \cdot N$ features via various feature extraction, where $\mathrm{Len}(dList)$ is the number of dilation rates and N is the number of features in a TSFresh set. In the prediction process, the same steps of feature extraction are used for feature extraction on unlabeled TS, as described in lines 17-26.

B. Ensemble Feature Selection

The aim of ensemble feature selection is to quickly obtain multiview features and enhance feature stability via integrating multiple filter feature selectors. The feature score is determined by measuring the correlation between features and labels, followed by generating multiview features utilizing different store ratios. In the training process, each feature score is calculated only once, and the top-ranked features are selected under different store ratios to serve as multiview features in lines 13-15. The efficient analysis of variance F-value is used to measure the correlation measure between features and labels. In the prediction process, multiview features are obtained directly using the trained feature indexes in line 28.

C. Multiview Ensemble classification

The multiview ensemble classification aims to integrate each view classifier's output fast to obtain the predicted label via hard voting. In the training process, the classifier is trained on each selected view feature vector, employing the ridge regression classifier with cross-validation in lines 14-16. In the prediction process, the predicted label is generated by integrating the classifiers' outputs on each view using hard voting in lines 27-29.

The limitations of the existing feature-based TSC algorithms in improving feature diversity and handling feature redundancy are alleviated to a certain extent through the reasonable coordination of various feature extraction, ensemble feature selection, and multiview ensemble classification processes of TD-MVDC.

TABLE I
CLASSIFICATION ACCURACY OF TD-MVDC AND OTHER FEATURE-BASED TSC ALGORITHMS

DataSets	1NN-DTW	Catch22	Signatures	TSFresh	FreshP	**TD-MVDC(ours)**
Adiac	60.36	71.61	72.12	81.07	**84.14**	79.54
ArrowHead	70.29	72.57	74.29	69.14	62.86	**81.14**
Beef	63.33	60.00	83.33	80.00	80.00	**96.67**
BeetleFly	70.00	75.00	90.00	50.00	90.00	**95.00**
BirdChicken	75.00	90.00	55.00	90.00	**100.00**	90.00
Car	73.33	76.67	76.67	83.33	76.67	**91.67**
CBF	**99.67**	96.67	94.22	99.56	99.44	98.67
ChlorineC	64.84	60.13	68.07	45.94	**77.47**	75.08
CinCECGTorso	65.07	81.74	90.94	**98.48**	96.67	95.00
Coffee	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
Computers	70.00	71.60	71.60	76.00	**76.80**	**76.80**
CricketX	**75.38**	58.21	62.05	69.74	71.03	66.92
CricketY	74.36	51.54	66.67	68.72	68.97	**78.72**
CricketZ	**75.38**	64.36	65.38	73.08	**75.38**	74.62
DiatomSizeR	**96.73**	94.44	86.60	30.07	85.29	94.77
DistalPOAG	**76.98**	70.50	74.82	73.38	75.54	75.54
DistalPOC	71.74	**79.71**	78.62	76.45	78.99	77.90
DistalPTW	58.99	66.19	66.91	64.03	69.06	**70.50**
Earthquakes	71.94	74.10	76.26	75.54	**76.98**	72.66
ECG200	77.00	84.00	83.00	85.00	**88.00**	**88.00**
ECG5000	92.44	93.93	93.91	94.20	94.42	**94.49**
ECGFiveDays	76.77	77.24	95.82	49.71	**100.00**	99.54
ElectricDevices	59.59	72.74	69.36	76.63	76.92	**77.98**
FaceAll	80.77	76.39	74.73	87.87	78.17	**93.67**
FaceFour	82.95	63.64	90.91	57.95	**96.59**	86.36
FacesUCR	90.49	70.73	83.37	87.12	89.76	**92.63**
FiftyWords	69.01	59.56	**74.51**	68.57	73.85	74.07
Fish	82.29	79.43	78.29	92.00	91.43	**97.71**
FordA	55.45	91.59	78.64	**100.00**	**100.00**	**100.00**
FordB	61.98	73.33	66.54	**83.09**	79.51	82.72
GunPoint	90.67	96.67	95.33	95.33	94.00	**99.33**
Ham	46.67	60.00	68.57	64.76	**74.29**	69.52
HandOutlines	88.11	85.95	89.46	90.27	**90.54**	**90.54**
Haptics	37.66	48.05	47.08	47.08	50.32	**56.49**
Herring	53.13	53.13	60.94	59.38	**65.63**	64.06
InlineSkate	38.36	42.55	32.36	35.64	**51.27**	48.36
InsectWingbeatS	35.51	56.72	65.20	65.86	65.71	**67.83**
ItalyPowerDemand	95.04	89.60	93.39	96.21	89.80	**96.50**
LargeKitchenA	79.47	84.53	61.87	81.60	**88.00**	83.20
Lightning2	**86.89**	68.85	70.49	54.10	75.41	65.57
Lightning7	72.60	69.86	72.60	65.75	71.23	**73.97**
Mallat	93.39	92.67	92.45	92.75	**96.38**	93.82
Meat	93.33	93.33	93.33	95.00	91.67	**96.67**
MedicalImages	73.68	74.87	72.50	78.03	**81.45**	75.66
MiddlePOAG	50.00	60.39	59.09	58.44	58.44	**66.23**
MiddlePOC	69.76	76.63	83.85	80.76	**86.94**	82.13
MiddlePTW	50.65	52.60	55.19	59.74	55.84	**60.39**
MoteStrain	83.47	86.58	90.10	53.91	**91.13**	**91.13**
NonIFECGT1	79.03	85.09	86.77	91.15	92.98	**93.33**
NonIFECGT2	86.46	87.28	89.77	93.54	93.64	**93.69**
OliveOil	83.33	73.33	**90.00**	86.67	86.67	**90.00**
OSULeaf	59.09	69.42	55.37	86.78	86.78	**97.11**
PhalangesOC	72.84	78.55	80.89	81.24	**83.80**	81.12
Phoneme	22.84	30.06	22.36	**36.66**	34.55	32.96
Plane	**100.00**	99.05	98.10	**100.00**	**100.00**	**100.00**
ProximalPOAG	80.49	84.39	84.88	84.39	83.41	**86.83**
ProximalPOC	78.35	83.51	87.63	86.94	**89.69**	89.00
ProximalPTW	75.61	76.10	**81.95**	78.05	77.56	78.05
RefrigerationD	46.40	52.27	54.40	56.00	**60.00**	55.73
ScreenType	39.73	50.93	49.07	**60.53**	58.40	47.73
ShapeletSim	65.00	98.89	65.56	50.00	**100.00**	96.67
ShapesAll	76.83	80.83	76.00	85.83	86.17	**93.50**
SmallKitchenA	64.27	80.80	**83.47**	81.33	81.60	82.93
SonyAIBORS1	72.55	85.69	72.38	91.68	91.85	**92.18**
SonyAIBORS2	83.11	92.65	82.16	90.24	91.61	**95.17**
StarLightCurves	90.66	96.95	95.64	97.95	97.97	**98.22**
Strawberry	94.05	92.70	95.68	94.86	95.95	**97.03**
SwedishLeaf	79.20	90.24	87.84	94.56	93.92	**97.76**
Symbols	94.97	96.08	91.86	93.67	97.29	**97.69**
SyntheticControl	99.33	96.33	98.67	**100.00**	99.67	99.33
ToeSegmentation1	77.19	**86.40**	67.98	52.63	81.58	85.53
ToeSegmentation2	83.85	79.23	82.31	83.85	86.15	**90.00**
Trace	**100.00**	**100.00**	98.00	**100.00**	**100.00**	**100.00**
TwoLeadECG	90.43	83.23	87.09	49.96	99.12	**99.56**
TwoPatterns	**100.00**	84.80	95.95	99.55	99.70	99.98
UWaveGestureLAll	89.17	82.80	94.72	95.11	96.37	**96.96**
UWaveGestureLX	72.75	75.46	80.37	80.35	84.03	**85.34**
UWaveGestureLY	63.40	69.54	71.41	72.53	**76.83**	75.35
UWaveGestureLZ	65.83	70.24	73.59	74.68	**78.59**	76.58
Wafer	97.99	99.74	99.21	**100.00**	**100.00**	**100.00**
Wine	57.41	38.89	77.78	50.00	79.63	**87.04**
WordSynonyms	64.89	53.76	59.25	58.78	**65.36**	60.03
Worms	58.44	72.73	62.34	76.62	**81.82**	74.03
WormsTwoClass	62.34	**85.71**	63.64	61.04	83.12	75.32
Yoga	83.63	78.40	80.73	**91.47**	**91.47**	89.77
Average ACC	74.02	76.38	77.31	76.82	83.29	**84.13**
Average rank	4.63	4.41	4.12	3.53	2.32	**1.99**
Performing best	10	5	5	11	32	**45**
Win	73	74	74	64	43	-
Draw	4	3	2	7	10	-
Lose	8	8	9	14	32	-

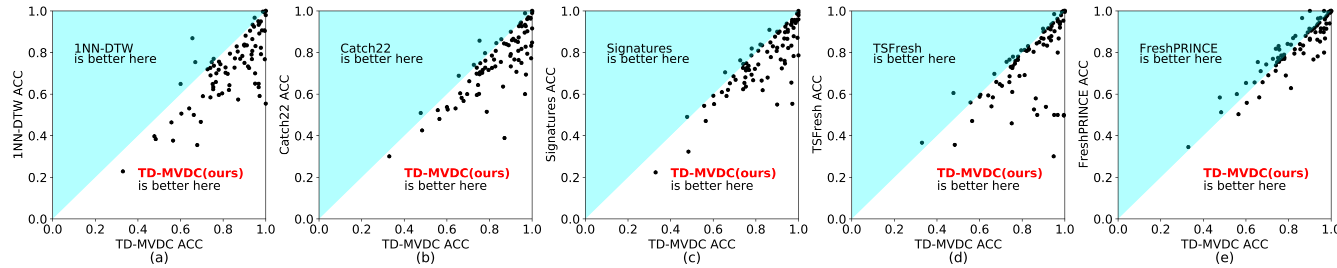

Fig. 3. Scatter plot of TD-MVDC and other TSC algorithms. (a) TD-MVDC versus 1NN-DTW. (b) TD-MVDC versus Catch22. (c) TD-MVDC versus Signatures. (d) TD-MVDC versus TS-Fresh. (e) TD-MVDC versus FreshPRINCE.

IV. Time Series Classification Experiment

A. Experiment Settings

To demonstrate the effectiveness of the proposed algorithm, the comparative and ablative classification experiments are conducted on 85 datasets sourced from the extensively employed UCR archive [19], where each dataset is divided into the default train/test set, and every sample is normalized. To evaluate the performance of the algorithms, the significance of the difference is measured using the Wilcoxon signed rank test with Holm correction at a p-value of 0.02, with accuracy being used as the performance index. All experiments are run on Python 3.8 and a laptop with AMD Ryzen 7 5800H CPU (8 cores, 3.20GHZ).

B. Comparative Experiment of Feature-based TSC Algorithms

In this section, the proposed TD-MVDC is compared with the representative feature-based TSC algorithms to demonstrate the performance advantages and characteristics. The contrast algorithms include the baseline distance-based DTW-1NN [8], and the feature-based Catch22 [11], Signatures [12], TSFresh [10], and FreshPRINCE [9], and their classification results are all from [8]. The classification accuracy of each algorithm is shown in Table I, where the last three rows represent the number of Win/Draw/Lose datasets when TD-MVDC is compared with other algorithms. The performing best row represents the number of datasets on which the algorithm achieved or tied for first place.

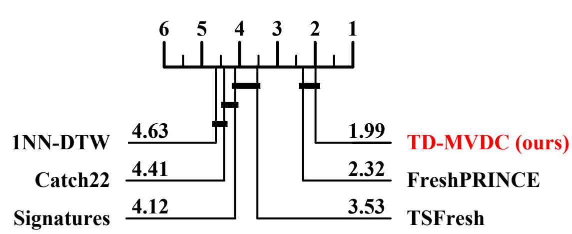

Fig. 4. Critical difference diagram of comparative experiment on 85 datasets of UCR archive.

The accuracy scatter plot Fig. 3 displays a detailed comparison, where points falling above/on/below the diagonal represent losses/draws/wins for TD-MVDC on that dataset. In every comparison, TD-MVDC wins on most datasets and has a significant accuracy lead on some datasets. Analysis of Table I and critical difference diagram Fig. 4 shows that TD-MVDC achieves the best overall performance among feature-based TSC algorithms, especially winning or tying for the championship on more than half (45/85) of the datasets. To obtain the applicable problem scope of TD-MVDC, the radar chart of TD-MVDC, TSFresh, and FreshPRINCE on UCR's seven category tasks is shown in Fig. 5. TD-MVDC performs best on four category tasks, particularly for Spectro and ECG, which provides a guide for its practical application. The main reason for failure on some issues is that TD-MVDC uses fixed filter factors to filter beneficial components, which is its limitation. The filter factor is fixed at 2, which compromises noise filtering and beneficial component retention, and its adaptive setting method will be studied in the future. In summary, TD-MVDC achieves state-of-the-art performance among feature-based TSC algorithms by improving feature diversity and stability and has the potential to handle Spectro and ECG applications.

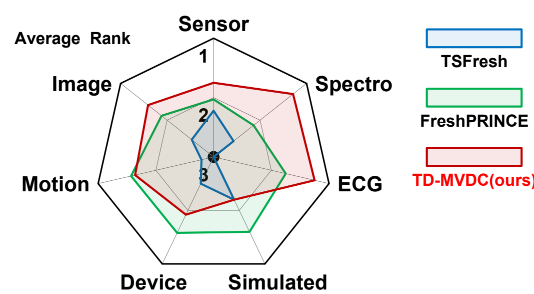

Fig. 5. Average rank radar chart of TSFresh, FreshPRINCE, and TD-MVDC on different types of UCR datasets.

C. Ablative Experiment

In this section, the proposed TD-MVDC is compared with its ablative versions to illustrate the contribution and importance of each component. The description of each ablation algorithm is shown in Table II, with FreshPRINCE as the baseline. The algorithm comparison results of the ablative experiment are shown in the critical difference diagram Fig. 6. The most prominent conclusion is that the performance of all ablative versions is lower than TD-MVDC, suggesting the indispensability of every component.

TABLE II
THE ALGORITHM DESCRIPTIONS FOR THE ABLATIVE VERSIONS OF TD-MVDC.

Ablative version	Description
TD-MVDC-A-Diff	The differential mapping is ablated.
TD-MVDC-A-TD	TD is replaced by the (1) to perform differential mapping.
TD-MVDC-A-Dila	The dilation mapping is ablated.
TD-MVDC-Whole	After the ensemble feature selection is ablated, all features extracted from various feature extractions are spliced into a whole feature vector.
View-0.1;0.2;0.3;0.4;0.5	After the hard voting is ablated, each view feature vector is obtained with different feature store ratios.

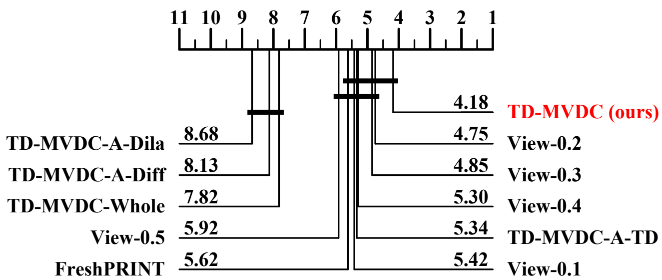

Fig. 6. Critical difference diagram of ablative experiment on 85 datasets of UCR archive.

The ablative versions of dilation mapping, differential mapping, and ensemble feature selection form groups and perform significantly worse than other algorithms, indicating that these are the most critical of the three components contributed by this article. Significantly, the worst performance is observed in the version with ablated dilation mapping, indicating that introducing the dilation mechanism is imperative for the future of feature-based TSC algorithms. It is also vital to use TD for differential transform, which shows that TD's filtering ability suppresses noise's negative impact on discriminative pattern recognition. Multiview versions with different feature store ratios have certain performance differences, and fast hard voting integrates the diverse prediction results between views to improve classification performance. In general, each component of TD-MVDC is essential, and feature diversity and feature stability make outstanding contributions to classification ability.

The results of TSC classification experiments show that TD-MVDC is a competitive feature-based TSC algorithm and is conceivable in industrial applications.

V. CONCLUSION

In this article, a novel feature-based TSC algorithm tracking differentiator-based multiview dilated characteristics (TD-MVDC) is proposed. Differential dilation mapping combines tracking differentiator and dilation mapping to suppress noise and improve the feature diversity of extracted TFresh features. Ensemble feature selection is based on filter feature selectors with different store ratios to enhance feature stability efficiently. Hard voting integrates the outputs of multiview classifiers to predict labels quickly. The classification performance of TD-MVDC is enhanced by the reasonable cooperation of each efficient component, as demonstrated by experiments conducted on the UCR archive. In future work, the adaptive parameter method of TD-MVDC will be studied and extended to multivariate TSC and practical tasks.

REFERENCES

[1] J. Li, Y. Shi, H. Li, *et al.*, "TC-GATN: temporal causal graph attention networks with nonlinear paradigm for multivariate time-series forecasting in industrial processes," *IEEE Trans. Ind. Informat.*, vol. 19, no. 6, pp. 7592-7601, Jun. 2023.

[2] X. Hu, L. Tan and T. Tang, "M^{2}BIST-SPNet: RUL prediction for railway signaling electromechanical devices," *J. Supercomput.*, 2024. https://doi.org/10.1007/s11227-024-06111-y

[3] J. Chae, S. Lee, J. Jang, *et al.*, "A survey and perspective on industrial cyber-physical systems (ICPS): from ICPS to AI-augmented ICPS," *IEEE Trans. Ind. Cyber- Phys. Syst.*, vol. 1, pp. 257–272, Oct. 2023.

[4] K. Leiner, F. P. Dollmann, M. F. Huber, *et al.*, "Cut interruption detection in the laser cutting process using ROCKET on audio signals," in *2023 IEEE 21st International Conference on Industrial Informatics (INDIN),* Jul. 2023, pp. 1–6.

[5] C. He, X. Huo, C. Zhu, *et al.*, "Feature selection-based multiview concentration for multivariate time series classification and its application," *IEEE Sens. J.*, vol. 24, no. 4, pp. 4798–4806, Feb. 2024.

[6] C. He, X. Huo, C. Zhu, *et al.*, "Minimum redundancy maximum relevancy-Based multiview generation for time series sensor data classification and its application," *IEEE Sens. J.*, vol. 24, no. 8, pp. 12830–12839, Apr. 2024.

[7] A. Bagnall, J. Lines, A. Bostrom, *et al.*, "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances," *Data Min. Knowl. Disc.*, vol. 31, no. 3, pp. 606-660, May. 2017.

[8] M. Middlehurst, P. Schafer and A. Bagnall, "Bake off redux: a review and experimental evaluation of recent time series classification algorithms," *Data Min. Knowl. Disc.*, 2024. https://doi.org/10.1007/s10618-024-01022-1

[9] M. Middlehurst and A. Bagnal, "The FreshPRINCE: A simple transformation based pipeline time series classifier," in *International Conference on Pattern Recognition and Artificial Intelligence,* May. 2020, pp. 150–161.

[10] M. Christ, N. Braun, J. Neuffer, *et al.*, "Time series feature extraction on basis of scalable hypothesis tests (tsfresh - a python package)," *Neurocomputing*, vol. 307, pp. 72–77, Sep. 2018.

[11] C. H. Lubba, S. S. Sethi, P. Knaute, *et al.*, "Catch22: canonical time-series characteristics selected through highly comparative time-series analysis," *Data Min. Knowl. Disc.*, vol. 33, no. 6, pp. 1821–1852, Nov. 2019.

[12] J. Morrill, A. Fermanian, P. Kidger, *et al.*, "A generalised signature method for multivariate time series feature extraction," *arXiv*, 2022, https://doi.org/10.48550/arXiv.2006.00873

[13] P. Schaefer and U. Leser, "WEASEL 2.0: a random dilated dictionary transform for fast, accurate and memory constrained time series classification," *Machine Learning*, vol. 112, no. 12, pp. 4763–4788, Dec. 2024.

[14] J. Han, "From PID to active disturbance rejection control," *IEEE Trans. Ind. Electron.*, vol. 56, no. 3, pp. 900–906, Mar. 2009.

[15] C. W. Tan, A. Dempster, C. Bergmeir, *et al.*, "MultiRocket: multiple pooling operators and transformations for fast and effective time series classification," *Data Min. Knowl. Disc.*, vol. 36, no. 5, pp. 1623–1646, Sep. 2022.

[16] V. Bolon-Canedo and A. Alonso-Betanzos, "Ensembles for feature selection: a review and future trends," *Information Fusion*, vol. 52, pp. 1–12, Dec. 2019.

[17] U. M. Khaire and R. Dhanalakshmi, "Stability of feature selection algorithm: a review," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1060–1073, Apr. 2022.

[18] B. Fulcher and N. Jones, "hctsa: A computational framework for automated time-series phenotyping using massive feature extraction," *Cell Systems.*, vol. 5, no. 5, pp. 527–531, Nov. 2017.

[19] H. A. Dau, A. Bagnall, K. Kamgar, *et al.*, "The UCR time series archive," *IEEE/CAA Journal of Automatica Sinica*, vol. 6, no. 6, pp. 1293–1305, Nov. 2019.