

# Time Series Classification with MuRBE: The Multiple Representation-Based Ensembles

**Rauzan Sumara**

*The Faculty of Mathematics and Information Science,*

*Warsaw University of Technology*

*Warsaw, Poland*

*rauzan.sumara.dokt@pw.edu.pl*

**Wladyslaw Homenda**

*The Faculty of Mathematics and Information Science,*

*Warsaw University of Technology, Warsaw, Poland*

*and*

*The Faculty of Applied Information Technology,*

*The University of Information Technology and*

*Management*

*Rzeszow, Poland*

*wladyslaw.homenda@pw.edu.pl*

**Witold Pedrycz**

*The Department of Electrical and Computer Engineering,*

*University of Alberta*

*Edmonton, Canada*

*wpedrycz@ualberta.ca*

**Fusheng Yu**

*The School of Mathematical Sciences,*

*Beijing Normal University*

*Beijing, China*

*yufusheng@bnu.edu.cn*

## Abstract

Time series classification has emerged as a pivotal endeavor in the realm of machine learning applications. This task is considered supervised learning, aimed at categorizing distinct classes within time series data. The present study introduces MuRBE (Multiple Representation-Based Ensembles), an innovative meta ensemble structure explicitly designed for time series classification. The MuRBE leverages the power of diverse representation domains, including feature-based, dictionary-based, interval-based, and shapelet-based methods. Exploiting complementary information from different representations makes it particularly effective to improve classification performance. A total of thirty distinguished benchmark datasets were utilized to evaluate the effectiveness of the proposed method, leading to competitive performance results. Notably, our approach secures a second rank among current state-of-the-art techniques.

**Keywords:** Machine learning, time series, classification, fuzzy-based ensemble, MuRBE.

## 1. Introduction

Within the realm of machine learning, time series classification has become a crucial task. As the complexity and volume of time series continue to grow, various algorithms have been developed. Recent investigations in ensemble techniques have received considerable interest due to their capacity to enhance accuracy by merging the advantages of multiple classifications, which served as the motivation behind our research.

In this study, we introduce MuRBE (Multiple Representation-Based Ensembles), a novel meta ensemble approach for time series classification. MuRBE leverages the power of diverse representation domains, including feature-based, dictionary-based, interval-

based, and shapelet-based methods. The study makes a unique contribution by integrating various domain classifiers through a fuzzy rank-based ensemble structure, which is not previously explored in the context of time series classification. The MuRBE aims to capture a wide range of temporal patterns and discriminatory characteristics. This approach leads to improved classification performance. By exploiting complementary information from different representations, making it particularly effective for complex time series classification tasks.

Our proposed MuRBE structure incorporates four different representation domains. Two of them are from recent studies, such as feature-based Autoregressive Fractional Integrated Moving Average with Random Forest (ARFIMA-RF) [12] and dictionary-based Symbolic Aggregate Approximation with Stacking Gated Recurrent Unit and Convolutional Neural Networks (SAX-SGCNN) [13]. We also take into consideration interval-based Diverse Representation Canonical Interval Forest (DrCIF) [8] and shapelet-based Random Dilated Shapelet Transform (RDST) [5].

We conducted an empirical experiment on 30 well-known benchmark datasets. We performed a comparative analysis, evaluating the classification outcomes of our approach against those achieved by current state-of-the-art methods. Furthermore, we have made the source code publicly available through a provided link <sup>1</sup>.

## 2. Literature Review

In contrast to conventional classification tasks wherein the sequence of attributes is irrelevant, time series classification entails the examination of temporally interrelated attributes, necessitating the analysis of comprehensive ordered sequences or time series data. The classification process involves predicting a class label for a sequence based on its measurable attributes or characteristic features. In turn, a classifier is utilized to distinguish between sequences that originate from different classes, with each sequence or time series having an equivalent set of extracted features.

A range of representation techniques has been specifically designed to classify time series data. Those techniques can be categorized based on the fundamental data representation employed. Feature-based approaches depend on global features extracted through a straightforward pipeline and fed into an appropriate classifier. Dictionary-based methods transform real-valued time series into discrete symbol sequences, thereby exploiting the frequency of recurrent patterns. Interval-based approaches generate features from specific time segments within the series, revealing temporal characteristics that might be obscured by irrelevant data. Shapelet-based approaches identify phase independent subsequences to effectively discriminate between time series.

The current state-of-the-art methods in the classification of time series is to utilize two or more representations. This methodology can be classified into four distinct categories. The first category is modular heterogeneous ensembles, where each component comprises a classifier designed based on a specific type of representation. For example, the Hierarchical Vote Collective of Transformation-based Ensembles (HIVE-COTE 2.0) or HC2 [8], at that time, was considerably more accurate on average compared to other established state-of-the-art techniques. Another approach, the Time Series Combination of Heterogeneous and Integrated Embedding Forest (TS-CHIEF) [11], is a classifier that bears the closest resemblance to HC2. TS-CHIEF consists of an ensemble of trees that incorporate distance, dictionary, and spectral-based features.

The second category comprises tree based homogeneous ensembles that incorporate a specific representation within the tree nodes. One notable technique is the Random Interval Spectral Ensemble (RISE) [4], an interval-based tree ensemble that employs a randomly unique chosen interval for every constituent classifier. In this method, the autoregression and periodogram functions are computed for each randomly selected interval, and these features are then combined into a feature vector. This vector is subsequently used to construct a tree. Another well-known approach is the Temporal Dictionary Ensemble (TDE) [8], which also falls under this category.

The third category is deep learning ensembles with embedded network representations.

<sup>1</sup> [github.com/rauzansumara/murbe-for-time-series-classification](https://github.com/rauzansumara/murbe-for-time-series-classification)

An example is InceptionTime [6], which combines five identical residual networks featuring inception modules. The fourth category utilizes convolution techniques to generate extensive new feature spaces, which are then analyzed using a linear classifier. One popular algorithm within the category is the ensemble of Random Convolutional Kernel Transform (ROCKET) models, also known as Arsenal [8]. The ROCKET [2] generates numerous summary statistics using randomly initialized convolutional kernels and then builds a linear ridge classifier to identify the classes.

It is also important to mention some of the current advanced methods that focus on a single representation. The Randomized Supervised Time Series Forest (RSTSf), introduced by [1], stands out as an interval-based tree incorporating a supervised technique for interval extraction, leveraging summary statistics and spectral features. FreshPRINCE [7] is a feature-based rotation forest classifier built under 800 extracted features from time series. Additionally, recent innovations include the Hybrid Dictionary-ROCKET Architecture (Hydra) [3] and Word Extraction for Time Series Classification with Dilation (WEASEL-D) [10]. These classifiers will also serve as benchmarks for comparison against the method proposed in this research.

### 3. The MuRBE Structure

This section provides a concise understanding of the MuRBE structure. The MuRBE is a heterogeneous ensemble having four modules each from a different representation. The component modules are: the ARFIMA-RF from feature-based representation [12]; SAX-SGCNN from the dictionary-based representation [13]; the interval-based DrCIF [8]; and the shapelet-based RDST [5]. These were chosen due to being one of the best in their domain representation. It is advisable to look at their original source to gain a comprehensive understanding of the base classifiers.

The four modules are ensembled using the fuzzy rank-based method, which was recently recognized as one of the most effective combinatorial approaches for different classifiers even with constrained domain expertise or antecedent knowledge [14]. Due to incorporating nonlinear functions to process decision scores, the fuzzy rank-based ensemble provides more flexible, dynamic, and adaptive weights of individual models based on their performance in specific contexts or regions of the input space. Unlike traditional ensemble methods, e.g., simple average or weighted average rules, they often use fixed weights, which may not adapt well to varying conditions.

The structure of MuRBE is presented in Figure 1. In the initial phase, each component is trained independently using standardized time series data and then required to produce a probability estimate (confidence score) for each class. After that, these probability scores from the base classifiers undergo transformation through two nonlinear functions: the exponential function and the hyperbolic tangent function. Due to their performance advantages, these nonlinear functions are commonly utilized to develop a tilted distribution, thereby emphasizing the distinctions between classifiers [14].

Figure 2 displays the graphs of these functions. The hyperbolic tangent function acts as a reward function while the exponential function acts as a decreasing function. The  $x$ -axis represents the probability of a class, where the exponential function measures the divergence from its objective for a class with a given prediction probability. As the input decreases, the divergence diminishes, ultimately reaching 0 when  $x$  equals 1. In contrast, the hyperbolic tangent function assesses the reward allocated to a class. The reward increases as the input grows, eventually reaching 1 when  $x$  equals 1. Consequently, using two nonlinear functions with varying concavities aims to yield complementary outcomes.

After mapping confidence scores from two functions with different concavities to create nonlinear fuzzy ranks, we combine these ranks to produce a rank score. This rank score is calculated as the product of divergence and reward corresponding to a specific confidence score. This procedure is repeated in each base classifier, and the rank scores are summed to determine the final fused score. A higher confidence score leads to a lower fused score, signifying a more accurate prediction. Therefore, the smallest fused score is considered the predicted class for the ensemble model. To enhance understanding of the concepts, we provide detailed steps for the proposed fuzzy rank-based ensemble as follows:

First, we calculate all the confidence scores. The confidence score of classes given by the base classifier is defined as  $p_k^i$ , where  $k = 1, 2, \dots, C$  is the number of classes and  $i = 1, 2, \dots, M$  is the number of base classifiers. Given that  $p_k^i \in [0, 1]$ , the probabilities  $(p_1^i, p_2^i, \dots, p_C^i)$  of  $C$  classes on the base classifier- $i$  essentially will satisfy the following condition,

$$\sum_{k=1}^C p_k^i = 1, \forall i = 1, 2, \dots, M. \quad (1)$$

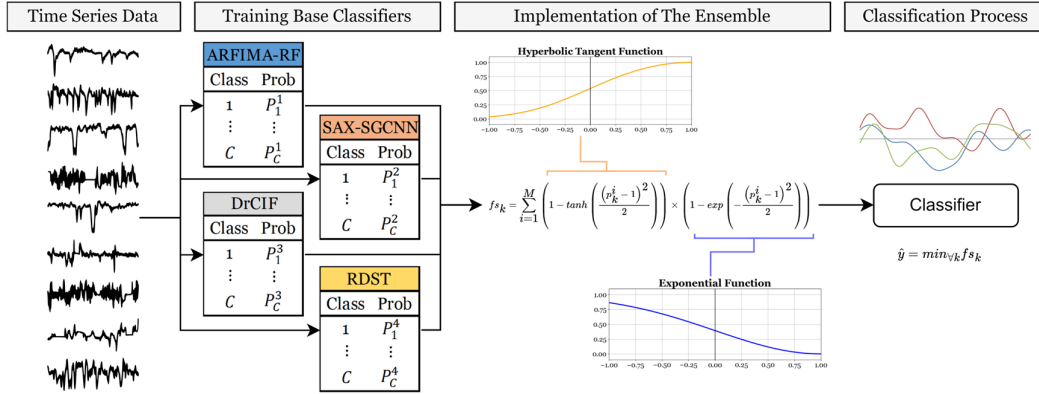


Fig. 1. The structure of the proposed method.

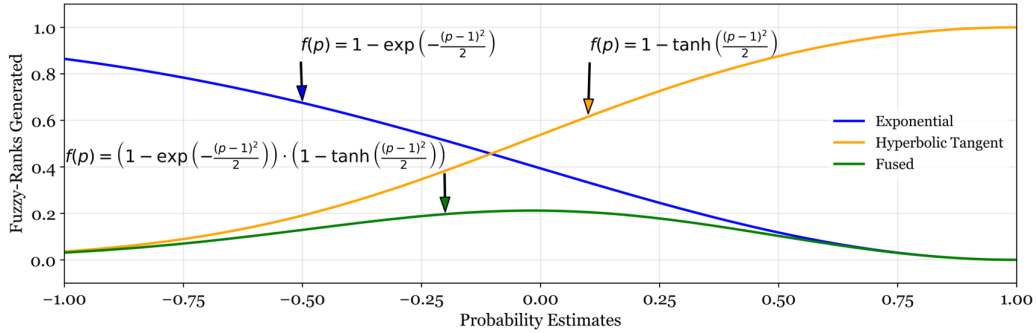


Fig. 2. Visualizations of nonlinear functions utilized for calculating fuzzy ranks.

Since we utilized two nonlinear functions, let us consider  $(r_1^{i1}, r_2^{i1}, \dots, r_C^{i1})$  and  $(r_1^{i2}, r_2^{i2}, \dots, r_C^{i2})$ , which are fuzzy ranks computed by using the hyperbolic tangent and the exponential functions expressed by

$$r_k^{i1} = 1 - \tanh\left(\frac{(p_k^i - 1)^2}{2}\right), \text{ and } r_k^{i2} = 1 - \exp\left(-\frac{(p_k^i - 1)^2}{2}\right). \quad (2)$$

Next, we can then determine the rank scores  $rs_k^i$  through the multiplication of fuzzy ranks  $r_k^{ij}$ . The fuzzy ranks are derived from nonlinear functions by substituting the confidence scores from the base classifiers. The standard notation for calculating the rank scores is presented as follows,

$$rs_k^i = \prod_{j=1}^S r_k^{ij}, \forall k = 1, 2, \dots, C \text{ and } \forall i = 1, 2, \dots, M, \quad (3)$$

where  $j = 1, 2, \dots, S$  is the number of chosen nonlinear functions, and the functions are bounded within  $[0, 1]$ . The product of the two nonlinear functions denoted as  $rs_k^i = r_k^{i1} \times r_k^{i2}$  is given in Figure 2. After that, the final fused score ( $fs_1, fs_2, \dots, fs_C$ ) is calculated by the equation as follows,

$$fs_k = \sum_{i=1}^M rs_k^i, \forall k = 1, 2, \dots, C. \quad (4)$$

The class with the minimum final fused score is selected as the predicted class. This fused score serves as the ultimate value for each class, which can be expressed through the following formula,

$$\hat{y} = \min_{\forall k} fs_k. \quad (5)$$

Figure 3 illustrates an example of how our proposed ensemble method works.

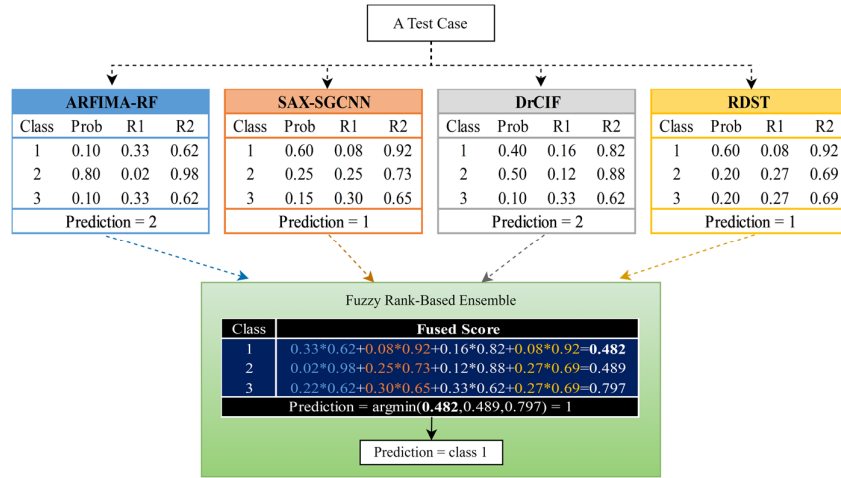


Fig. 3. An overview of the proposed ensemble structure for a three-class problem.

#### 4. Experimental Results

We conducted experiments using 30 out of 112 publicly available time series datasets from the UCR/UEA time series classification archive [15]. The archive is commonly used in data mining and machine learning literature. We proportionally selected these datasets according to the number of accessible datasets in each category to ensure a comprehensive evaluation. This strategy improved the robustness of the study by enabling us to maintain a balanced representation across several categories.

The selected datasets had diverse properties, such as different numbers of classes and lengths of series, and different numbers of observations given on the train and test sets. The type of datasets also differs. We have Spectro data (Coffee, Meat, OliveOil), Simulated data (ShapeletSim, SyntheticControl, TwoPatterns), Sensor data (Earthquakes, FordA, Lightning7, Trace), Motion data (CricketZ, GunPoint, InlineSkate), Image data (BeetleFly, BirdChicken, FaceFour, FiftyWords), and more. A comprehensive evaluation was possible due to these differences. The datasets are presented in Table 1. Due to space limitations, we cannot include a full summary here. However, it is available on the cited website.

The experiments were performed using Python 3.11.5 on a computer with an 8-core CPU, 16 gigabytes of RAM, GeForce GTX 1660 Ti graphics card (GPU), and the Windows 11 operating system. Subsequently, the train set from the repository was operated to train the models, given that the datasets within the repository were already divided into train and

test sets. We utilize the same train-test split as is to ensure our results are comparable to the previous study. During the evaluation, the compared results are limited to the test sets only. To maintain consistency with existing literature, we used the default parameter configuration setting described in its literature for the base classifiers in MuRBE. More complete default setting reviews can also be found in [9]. For the state-of-the-art methods used as competitors in the study, the default settings for the parameter configuration were applied. Last but not least, we also provided critical difference diagrams (CD) as a post-hoc analysis, employing the Wilcoxon-Holm method with  $\alpha = 5\%$ .

Table 1 only provides the associated accuracy of base classifiers and MuRBE. Average accuracy and rank are also offered in the last two rows of the table, respectively. As a result, the MuRBE consistently outperforms the base classifiers with an average accuracy of over 1% more accurate. Our core result is that the proposed method proves to improve classification performance in terms of accuracy. Moreover, the CD diagram in Figure 4(a) displays MuRBE compared to the base classifiers based on the average ranks on 30 datasets. A lower average rank implies better accuracy of a certain method, and solid bars group classifiers between two or more methods for which there are statistically insignificant differences. Out of 30 datasets, The MuRBE ranks one above all the base classifiers, demonstrating its superiority over four base classifiers.

**Table 1.** Our proposed MuRBE compared to base classifiers in terms of accuracy.

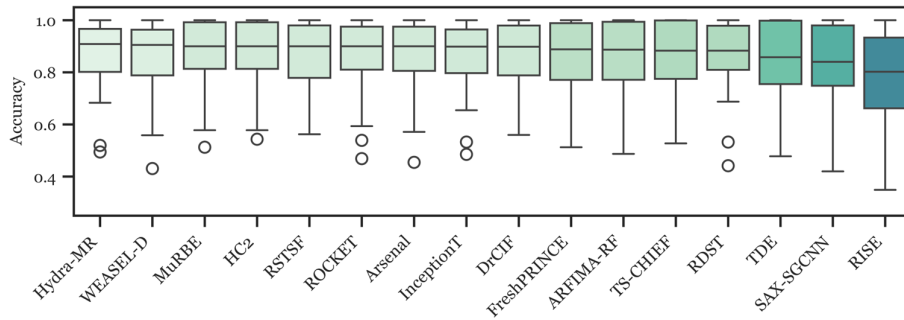
| Datasets                      | ARFIMA-RF     | SAX-SGCNN     | DrCIF         | RDST          | MuRBE         |
|-------------------------------|---------------|---------------|---------------|---------------|---------------|
| BeetleFly                     | 0.9000        | 1.0000        | 0.9000        | 0.9500        | 0.9000        |
| BirdChicken                   | 1.0000        | 1.0000        | 0.9500        | 0.9000        | 0.9500        |
| Coffee                        | 1.0000        | 1.0000        | 1.0000        | 1.0000        | 1.0000        |
| CricketZ                      | 0.7487        | 0.7870        | 0.8026        | 0.8564        | 0.8590        |
| DistalPhalanxOutlineCorrect   | 0.8007        | 0.7790        | 0.7826        | 0.7754        | 0.7754        |
| DistalPhalanxTW               | 0.6835        | 0.7754        | 0.6906        | 0.7122        | 0.7194        |
| Earthquakes                   | 0.7482        | 0.7681        | 0.7482        | 0.7266        | 0.7482        |
| ECG200                        | 0.8600        | 0.8200        | 0.8800        | 0.8900        | 0.8600        |
| FaceFour                      | 0.9659        | 0.9545        | 0.9886        | 0.9886        | 1.0000        |
| FiftyWords                    | 0.7231        | 0.9890        | 0.7978        | 0.8527        | 0.8330        |
| FordA                         | 1.0000        | 0.8974        | 0.9682        | 0.9447        | 0.9561        |
| GunPoint                      | 0.9800        | 1.0000        | 0.9933        | 1.0000        | 1.0000        |
| Herring                       | 0.5625        | 0.7094        | 0.6406        | 0.6875        | 0.6094        |
| InlineSkate                   | 0.4873        | 0.4202        | 0.5600        | 0.4418        | 0.5436        |
| ItalyPowerDemand              | 0.9572        | 0.7467        | 0.9689        | 0.9397        | 0.9699        |
| LargeKitchenAppliances        | 0.8747        | 0.8427        | 0.8240        | 0.8267        | 0.9040        |
| Lightning7                    | 0.7671        | 0.7397        | 0.7534        | 0.8082        | 0.8082        |
| Meat                          | 0.9333        | 0.9267        | 0.9500        | 0.9333        | 0.9333        |
| MedicalImages                 | 0.7855        | 0.7276        | 0.7882        | 0.7645        | 0.8066        |
| MiddlePhalanxOutlineCorrect   | 0.8557        | 0.7553        | 0.8316        | 0.8419        | 0.8488        |
| MiddlePhalanxTW               | 0.6039        | 0.6202        | 0.5909        | 0.5325        | 0.5779        |
| MoteStrain                    | 0.9529        | 0.9233        | 0.9353        | 0.9337        | 0.9588        |
| OliveOil                      | 0.9000        | 0.9193        | 0.9333        | 0.8667        | 0.8667        |
| Plane                         | 1.0000        | 1.0000        | 1.0000        | 1.0000        | 1.0000        |
| ProximalPhalanxOutlineCorrect | 0.8591        | 0.8385        | 0.8969        | 0.8763        | 0.9003        |
| ProximalPhalanxTW             | 0.8000        | 0.7260        | 0.7902        | 0.8146        | 0.8293        |
| ShapeletSim                   | 1.0000        | 0.9889        | 0.9833        | 0.9889        | 1.0000        |
| SyntheticControl              | 1.0000        | 0.6233        | 1.0000        | 0.9933        | 1.0000        |
| Trace                         | 1.0000        | 0.8800        | 1.0000        | 1.0000        | 1.0000        |
| TwoPatterns                   | 0.9988        | 0.9963        | 0.9998        | 1.0000        | 1.0000        |
| <b>Average accuracy</b>       | <b>0.8583</b> | <b>0.8385</b> | <b>0.8649</b> | <b>0.8615</b> | <b>0.8719</b> |
| <b>Average rank</b>           | <b>3.0667</b> | <b>3.5500</b> | <b>2.8833</b> | <b>3.1167</b> | <b>2.3833</b> |

Apart from that, we also consider several competitors mentioned in the literature. They are under the current state-of-the-art methods, such as ROCKET, Arsenal, RSTSF, HydraMR, FreshPRINCE, WEASEL-D, InceptionTime, TS-CHIEF, RISE, TDE, and HC2.

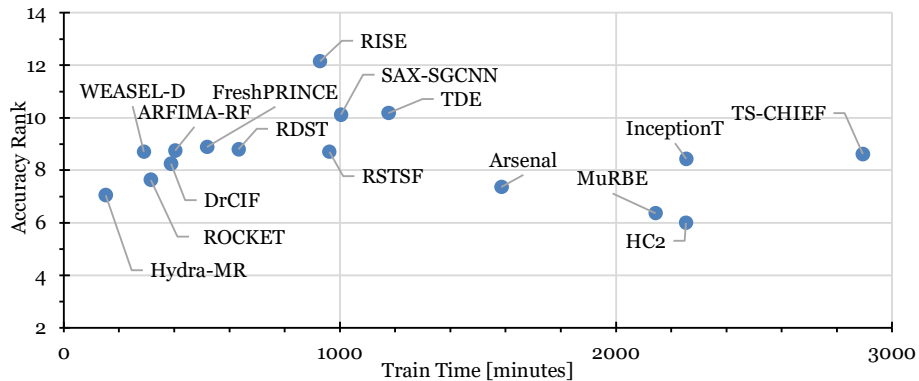
Figure 4(b) illustrates the average ranks of MuRBE against each of the leading state-of-the-art algorithms on 30 datasets. Overall, the proposed MuRBE method is highly competitive compared to the other methods. In particular, it outperforms the most current methods except HC2 in terms of accuracy. Based on average ranks, MuRBE achieves the second position and exhibits a slight edge over the leading fast algorithms such as Hydra-MR, ROCKET, and WEASEL-D. It faintly underperforms as opposed to HC2 as the toughest and the most accurate competitor in the group, even though their differences are statistically insignificant. Our findings also indicate that the proposed MuRBE achieves more pairwise wins in Image, Simulated, Electrocardiogram (ECG), and Spectro data types.



**Fig. 4.** CD diagram based on the average rank of MuRBE in comparison with (a) the base classifiers and (b) other competitors.



**Fig. 5.** A box plot representing test accuracy for each algorithm.



**Fig. 6.** Examining accuracy rank and train time over 30 datasets.

Nevertheless, we conclude that the dataset type strongly affects the classification quality. In this regard, we would like to emphasize that none of the methods can outperform the other methods in all cases. Boxplots displayed in Figure 5, arranged by descending medians across

algorithms, illustrate how close the accuracy of MuRBE is in comparison with the current state-of-the-art methods. Top-ranking algorithms have a narrow interquartile range (IQR), few outliers, and a median accuracy exceeding 80%. The boxplot also shows that Hydra-MR and WEASEL-D achieve the highest median among other algorithms. However, their medians tend to exceed the average accuracy, indicating a strong left skew. It signifies the presence of strong extremely low accuracy values, i.e., outliers, which stretch the boxplot distribution to the left. In contrast, the HC2 and MuRBE have an average accuracy nearly identical to their medians, indicating a weak skewness in their distributions. Moreover, the boxplot distributions of the proposed MuRBE and HC2 are related. This relation might be due to the fact that both methods use DrCIF as one of their component classifiers within their ensemble structures. As illustrated in Figure 4(a), DrCIF is a particularly strong base classifier compared to the others. Nevertheless, a more detailed explanation of the reasons behind this comparison will be provided in our future research.

Although we believe that accuracy is crucial, we also provide the runtime as another factor in assessing algorithm performance. Figure 6 depicts the average accuracy rank against the runtime for all compared algorithms, revealing a clear trade-off between performance and runtime. Several important caveats should be noted when interpreting these results. Firstly, all algorithms, with the exception of InceptionTime, were executed on a single-thread CPU. Therefore, the runtime associated with InceptionTime is not directly comparable due to its operation on a GPU. Additionally, the maximum memory usage was not explicitly measured in this experiment, but it remained within the 16GB limit of the available memory on the CPU. The reported running times are sequential, reflecting the sequential memory usage. It is important to note that if all algorithms were multithreaded, the runtime would significantly decrease, of course.

Considering these factors, we can draw several conclusions. Hydra-MR and ROCKET are capable of training models for all 30 datasets in less than 3 hours and 6 hours, respectively, even without multithreading. If time efficiency is the primary concern, Hydra-MR or ROCKET would be an excellent starting point for any analysis. On the other hand, TS-CHIEF is significantly slower and appears to scale less well than the other algorithms. Additionally, we see SAX-SGCNN as the slowest component within MuRBE, although this is attributed to the configuration of hyperparameter settings rather than inherent limitations. Specifically, it involved a wide range of hyperparameter tuning, namely the alphabet size and word length. For a small dataset, it requires much more runtime than other components (even though it is still under an hour). This will be an area for future improvement. Meanwhile, HC2 or MuRBE can be a worthy choice when prioritizing model accuracy over runtime.

## 5. Conclusion

This paper presents MuRBE as a novel meta ensemble approach for time series classification, which is designed from four different representation domains. Among the state-of-the-art algorithms, our proposed MuRBE outperforms them except for HC2. We are convinced its main advantage can be attributed to the presence of discriminatory features in multiple representation domains, which are common in many datasets. However, a significant limitation lies in its computational demands. Training time can become excessively long for large-scale problems involving tens of thousands of series with lengths in the tens of thousands or even more. The method might not be well scalable for such issues, but we believe there is potential for improvement. Future research could focus on enhancing individual components, particularly SAX-SGCNN, to enable more adaptive and intelligent reconfiguration settings for optimal runtime. Developing better strategies to address the issue will also be a key focus in future modifications.

## Acknowledgments

This project received funding from a grant provided by the Faculty of Mathematics and Information Science at the Warsaw University of Technology.

## References

1. Cabello, N., Naghizade, E., Qi, J., Kulik, L.: Fast, Accurate and Interpretable Time Series Classification Through Randomization. (2021)
2. Dempster, A., Petitjean, F., Webb, G.I.: ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Min Knowl Disc.* 34 (5), 1454–1495 (2020)
3. Dempster, A., Schmidt, D.F., Webb, G.I.: HYDRA: Competing convolutional kernels for fast and accurate time series classification, <http://arxiv.org/abs/2203.13652>, (2022)
4. Flynn, M., Large, J., Bagnall, T.: The Contract Random Interval Spectral Ensemble (c-RISE): The Effect of Contracting a Classifier on Accuracy. In: Pérez García, H., Sánchez González, L., Castejón Limas, M., Quintián Pardo, H., and Corchado Rodríguez, E. (eds.) *Hybrid Artificial Intelligent Systems*. pp. 381–392. Springer International Publishing, Cham (2019)
5. Guillaume, A., Vrain, C., Elloumi, W.: Random Dilated Shapelet Transform: A New Approach for Time Series Shapelets. In: El Yacoubi, M., Granger, E., Yuen, P.C., Pal, U., and Vincent, N. (eds.) *Pattern Recognition and Artificial Intelligence*. pp. 653–664. Springer International Publishing, Cham (2022)
6. Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D.F., Weber, J., Webb, G.I., Idoumghar, L., Muller, P.-A., Petitjean, F.: InceptionTime: Finding AlexNet for time series classification. *Data Min Knowl Disc.* 34 (6), 1936–1962 (2020)
7. Middlehurst, M., Bagnall, A.: The FreshPRINCE: A Simple Transformation Based Pipeline Time Series Classifier. (2022)
8. Middlehurst, M., Large, J., Flynn, M., Lines, J., Bostrom, A., Bagnall, A.: HIVE-COTE 2.0: a new meta ensemble for time series classification. *Mach Learn.* 110 (11–12), 3211–3243 (2021)
9. Middlehurst, M., Schäfer, P., Bagnall, A.: Bake off redux: a review and experimental evaluation of recent time series classification algorithms. *Data Min Knowl Disc.* 38 (4), 1958–2031 (2024)
10. Schäfer, P., Leser, U.: WEASEL 2.0 -- A Random Dilated Dictionary Transform for Fast, Accurate and Memory Constrained Time Series Classification. (2023)
11. Shifaz, A., Pelletier, C., Petitjean, F., Webb, G.I.: TS-CHIEF: a scalable and accurate forest algorithm for time series classification. *Data Min Knowl Disc.* 34 (3), 742–775 (2020)
12. Sumara, R., Homenda, W., Pedrycz, W.: ARFIMA for Feature-Based Time Series Classification. *PACIS 2024 Proceedings*. (2024)
13. Sumara, R., Homenda, W., Pedrycz, W., Yu, F.: A Dictionary-Based with Stacked Ensemble Learning to Time Series Classification. In: Franco, L., De Mulatier, C., Paszynski, M., Krzhizhanovskaya, V.V., Dongarra, J.J., and Sloot, P.M.A. (eds.) *Computational Science – ICCS 2024*. pp. 120–128. Springer Nature Switzerland, Cham (2024)
14. Valenkova, D., Lyanova, A., Sinitca, A., Sarkar, R., Kaplun, D.: A fuzzy rank-based ensemble of CNN models for MRI segmentation. *Biomedical Signal Processing and Control*. 102 107342 (2025)
15. Time Series Classification Website, <https://www.timeseriesclassification.com/>, Accessed: January 02, 2025