Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 10 No. 5 (2026) 1233 - 1242 e-ISSN: 2580-0760

A Hybrid Dynamic Time Warping and Multi-BMU Self-Organizing Map for Time-series Clustering

Rina Dewi Indahsari[1], Siti Sendari[2], Wahyu Sakti Gunawan Irianto[3], Samsul Arifin[4], Dyah Lestari[5]
[1,2,3,5]Departement of Electronics and Informatics Engineering, Universitas Negeri Malang, Malang, East Java, Indonesia
[4]Departement of Computer System, Institut Teknologi dan Bisnis Asia Malang, Malang, East Java, Indonesia

*rina.dewi.2205349@students.um.ac.id

Abstract

Time-series clustering remains challenging because many SOM-based studies still improve either temporal similarity measurement or winner adaptation separately, leaving limited evidence on the effect of integrating both components in a single framework. This study proposes a hybrid Self-Organizing Map (SOM) that combines Dynamic Time Warping (DTW) and a multi-Best Matching Unit (multi-BMU) mechanism for temporal pattern clustering. The proposed approach replaces Euclidean distance with DTW in representative-neuron selection and extends learning from a single BMU to a top-k winner scheme with k=3. The model was evaluated on 2,505 room-lighting time-series samples of length 25 collected over 60 weeks in an office building in Indonesia. Four models were compared across 10 independent runs: M1 (standard SOM), M2 (SOM + DTW), M3 (SOM + multi-BMU), and M4 (hybrid SOM + DTW + multi-BMU). The results show that M4 achieved the best clustering quality, with the highest mean Silhouette score (0.768300), the lowest Davies–Bouldin Index (0.4011), and the highest Calinski–Harabasz Index (233.1750). In contrast, M1 retained the lowest mean Topographic Error (0.150858). Statistical tests confirmed significant differences among models. Although the proposed model increases computational complexity compared with standard SOM, the use of a DTW window constraint helps reduce the alignment cost while preserving temporal sensitivity. The integration of DTW and multi-BMU substantially improves temporal cluster quality, although it is accompanied by reduced topological preservation and increased computational complexity.

Keywords: dynamic time warping; multi-BMU; self-organizing map; temporal pattern analysis; time-series clustering

How to Cite: R. D. Indahsari, S. Sendari, W. S. . Gunawan Irianto, S. . Arifin, and D. Lestari, "A Hybrid Dynamic Time Warping and Multi-BMU Self-Organizing Map for Time-series Clustering", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 10, no. 5, pp. 1233–1242, Sep. 2026, doi: 10.29207/resti.v10i5.7937.
Permalink/DOI: https://doi.org/10.29207/resti.v10i5.7937

Received: May 26, 2026
Accepted: September 5, 2026
Available Online: September 29, 2026

This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Time-series data have become increasingly important as digital technologies continue to expand data generation across many domains [1-3]. Such data capture real-world phenomena as sequences of observations that evolve over time [4], [5]. In practice, time-series data are widely produced in strategic fields such as healthcare, meteorology, environmental monitoring, industry, and agriculture [2], [6], [7]. Consequently, temporal data analysis is no longer viewed merely as the processing of ordered numerical observations, but as an effort to understand complex, nonlinear, and often irregular system dynamics across time [8], [9]. For this reason, temporal pattern extraction and time-series clustering have become essential for descriptive analysis, exploratory understanding, and data-driven decision-making [4], [10], [11].

Recent advances in time-series clustering show substantial variation in similarity measures, representation strategies, and learning algorithms[8], [12], [13]. In many studies, conventional distance measures, especially Euclidean distance, remain widely used because of their computational simplicity and ease of integration into clustering algorithms [12], [14]. However, for temporal data, direct point-to-point comparison is often unrealistic because two series that represent the same underlying pattern may appear with phase shifts, different rates of evolution, or local temporal distortions [4], [15]. Therefore, for time-series analysis, similarity should be defined not only by value proximity at identical timestamps, but by the correspondence of the overall shape and temporal

dynamics of the series [8], [16-18]. One of the most influential approaches for handling this issue is Dynamic Time Warping (DTW), which allows nonlinear alignment between two temporal sequences so that similar patterns can still be identified even when they are not temporally synchronized [17], [19], [20]. Although DTW is more appropriate for temporal data because it can accommodate time shifts and local distortions, its computational cost is substantially higher than that of Euclidean distance, and its integration into clustering frameworks is not always straightforward [4], [21], [22]. As a result, recent studies have moved toward hybrid approaches that combine alignment-aware similarity with unsupervised learning mechanisms that remain interpretable, stable, and computationally feasible [4], [15], [19].

Self-Organizing Maps (SOMs) remain one of the most widely used clustering methods because they combine clustering, dimensionality reduction, and topological preservation within a single unsupervised learning framework [6], [23], [24]. Unlike many other clustering algorithms, SOMs not only assign data to clusters but also map them into a lower-dimensional space, allowing relationships among patterns and latent data structures to be inspected visual [4], [10], [25], [26]. SOMs are therefore valuable not only for cluster formation, but also for exploration and interpretation of complex high-dimensional data [6], [23], [27]. However, conventional SOMs still face fundamental limitations when applied to time-series data, particularly because the Best Matching Unit (BMU) is determined using Euclidean distance. This makes the model sensitive to temporal misalignment and less effective at capturing similarity when the same pattern occurs at different temporal positions [8], [19]. In addition, conventional SOM learning depends heavily on a single winning neuron, which makes prototype updating overly concentrated and may neglect other neurons that are substantively close to the input pattern [28]. Its performance is also influenced by weight initialization, noise sensitivity, map stability, and the risk of weak representations when the data distribution is complex or heterogeneous [23], [28]. These limitations indicate that the original SOM mechanism is not yet fully adequate for temporal clustering tasks that require sensitivity to sequence dynamics together with more robust representations [4], [29].

Several previous studies have attempted to address these limitations through different forms of SOM modification. Principal Component Analysis (PCA)-based initialization has been shown to accelerate convergence and improve map stability [30], [28]. DTW-SOM by [19] showed that incorporating DTW into the matching process improves the recognition of temporally distorted but structurally similar patterns. [6] introduced a structural SOM (S-SOM) that emphasizes structural similarity in representative neuron search. [8] modified SOM for time-series clustering and reported better performance than conventional hierarchical approaches. [4] proposed SOMTimes, which combines DTW with a pruning strategy to improve temporal processing efficiency. Meanwhile, [28] introduced vwSOM, showing that virtual winning neurons can enrich the learning process, whereas [23] demonstrated that randomization can improve topological map organization.

Nevertheless, these improvements have largely been developed in isolation. Some studies focus primarily on alignment-aware similarity through DTW [4], [19], whereas others emphasize extensions of the winner-update mechanism [28]. However, time-series clustering requires both: the ability to recognize patterns that are similar despite temporal misalignment and the ability to form prototype representations that are not overly dependent on a single winning neuron. This suggests the need for a more integrated SOM framework that combines temporal alignment and distributed prototype learning. Thus, the main research gap addressed in this study is the limited number of studies that explicitly and systematically integrate DTW as a similarity measure with a multi-BMU mechanism as a learning strategy in SOM for time-series clustering [4], [19], [28]. This gap is important because the available empirical and conceptual evidence regarding whether the combination of alignment-aware similarity and distributed winner adaptation can jointly improve temporal clustering quality remains limited [8], [28].

In addition, existing SOM-based time-series clustering studies have generally improved either temporal similarity modeling or winner adaptation separately, leaving insufficient evidence on whether both mechanisms can be integrated effectively within a unified SOM learning framework. Therefore, the present study is positioned not merely as another application of DTW in SOM, nor as a simple extension of multi-winner learning, but as an attempt to combine both components in a coherent architecture that is specifically designed for temporal clustering.

Based on this gap, the present study aims to develop a hybrid DTW and multi-BMU approach within the SOM algorithm for time-series data. Specifically, the study integrates DTW into representative unit identification so that similarity becomes more sensitive to temporal shape correspondence, while multi-BMU learning allows weight updates to involve several nearby neurons that jointly represent the input pattern. The main contribution of this study is the Equationtion of a SOM learning framework that combines alignment-aware similarity and distributed winner adaptation to improve temporal sensitivity, enrich prototype representation, and strengthen clustering quality.

To address this objective, the study is guided by the following research questions:

RQ1: Does the integration of DTW improve SOM clustering quality for time-series data compared with standard SOM?

RQ2: Does the multi-BMU mechanism improve prototype representation and clustering quality compared with the conventional single-BMU SOM?

RQ3: Does the combination of DTW and multi-BMU produce a statistically significant improvement over standard SOM and its partial variants?

2. Methods

This study employed a quantitative-computational approach designed as an experimental computational study to develop and evaluate a clustering model for time-series data. Operationally, the study comprised four stages: data preparation, including acquisition, cleaning, sequence-length alignment, and normalization; model construction, including implementation of conventional SOM as the baseline and development of the hybrid SOM-DTW-multi-BMU model; comparative experimentation through an ablation study; and performance evaluation using the Silhouette score and Topographic Error (TE), together with across-run stability analysis and statistical testing. The research stages are shown in Figure 1.

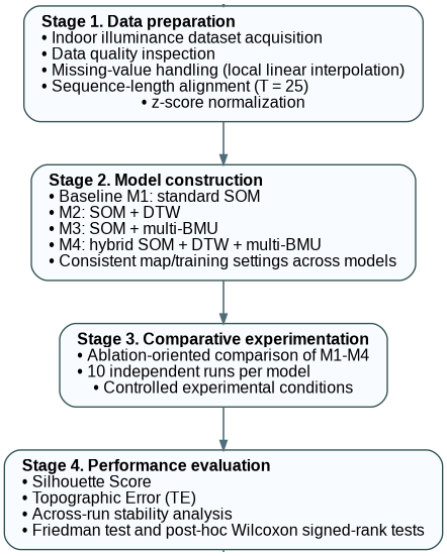


Figure 1. Research Workflow

2.1 Dataset and Preprocessing

The data used in this study were primary indoor illuminance measurements recorded over 60 weeks in an office building in Indonesia. The dataset consists of 2,505 samples, each with a time-series length of 25. Each series represents the temporal illuminance profile of one observation unit, so that the full dataset forms a set of temporal patterns reflecting variation in space-use behavior, light intensity, building operational rhythms, and possible fluctuations influenced by environmental conditions and occupant activities. The dataset was selected for both substantive and methodological reasons. Substantively, indoor lighting patterns constitute a real temporal phenomenon that is not fully stationary and may display periodicity, local change, and similar shapes that are not perfectly synchronized over time. Methodologically, these characteristics make the dataset suitable for evaluating clustering models that are sensitive to shape similarity and temporal distortion. Moreover, the relatively large sample size supports a more robust evaluation for unsupervised time-series clustering, whereas the moderate sequence length preserves temporal complexity without imposing excessive computational burden.

Data preprocessing included data quality inspection, time-series length alignment, z-score normalization, and light noise handling. Data quality inspection was performed to identify missing values, unreasonable extreme values, and sensor recording errors. Missing values occurring in small proportions were handled through local linear interpolation in order to preserve temporal continuity. Z-score normalization mapped each time series onto a comparable scale, enabling the model to focus more strongly on curve shape, relative variation, and temporal dynamics.

2.2 Conventional SOM as Baseline

Conventional SOM was used as the main baseline. Let the dataset be denoted as $X = \{x_1, x_2, \dots, x_n\}$ with $x_i \in R^T$, T=25, and the SOM map consisting of M neurons arranged on a two-dimensional lattice. Each neuron j has a weight vector $w_j \in R^T$. For each input x, the representative neuron is determined as the BMU based on Euclidean distance in Equation 1 [28].

$$c = \arg min_j \|x - w_j\|_2 \quad (1)$$

After the BMU c is obtained, the neuron weights are updated using a Gaussian neighborhood function according to Equation 2 [29].

$$w_j(t+1) = w_j(t) + \alpha(t) h_{c,j}(t)(x - w_j(t)) \quad (2)$$

$\alpha(t)$ is the learning rate at iteration t and $h_{c,j}(t)$ is the neighborhood function (Equation 3) [29].

$$h_{c,j}(t) = exp\left(-\frac{\|r_c - r_j\|^2}{2\sigma(t)^2}\right) \quad (3)$$

r_c and r_j denote neuron positions on the lattice, and $\sigma(t)$ is the neighborhood radius that decreases throughout training. The learning rate and radius decay exponentially as given in Equation 4 [28].

$$\alpha(t) = \alpha_0 exp\left(-\frac{t}{\tau_\alpha}\right), \quad \sigma(t) = \sigma_0 exp\left(-\frac{t}{\tau_\sigma}\right) \quad (4)$$

Although this structure is effective for static vector data, it is less suitable for time-series data because Euclidean-based BMU search cannot handle temporal misalignment, while the single-winner scheme tends to produce overly narrow representations.

2.3 SOM Training Parameters

The performance of the SOM is influenced by several training parameters, including grid size, learning rate, neighborhood radius, number of iterations, and the number of winning neurons. In this study, the SOM grid was set to 15 × 15, resulting in 225 neurons. This configuration follows the principle proposed by

Vesanto and Alhoniemi [31], namely $M = 5\sqrt{\text{N}}$, where M denotes the number of map units and N is the number of samples. This grid size was selected to provide sufficient representational detail while maintaining map readability.

The initial learning rate was set to 0.5 and gradually decreased during training. This setting allows the model to learn the global data structure more quickly in the early phase and then converge more smoothly in later iterations. The neighborhood radius was set to 1.0, which determines the extent to which neighboring neurons are affected during weight updating. This value was considered adequate to preserve a balance between global organization and local adaptation.

The number of training iterations was set to 2000 in order to ensure that the map reached a sufficiently stable state without excessive computational cost. In addition, the proposed model employs an N-BMU mechanism with k=3, meaning that the three most similar neurons contribute to the update process. This value was chosen to enrich prototype adaptation while avoiding an overly diffused learning process. Overall, these parameter settings were designed to produce a stable and representative SOM for temporal pattern analysis.

2.4 PCA-based Weight Initialization

To improve the stability of early convergence, the proposed model uses PCA-based weight initialization. The initial vectors are projected along the first two principal components of the normalized data. This strategy was selected because PCA provides a global orientation of the data, allowing the initial neuron distribution to better reflect the dominant structure of the dataset. In the SOM context, PCA initialization helps accelerate map organization and reduces between-run performance variation caused by random initialization. If $\bar{x}$ denotes the data mean and p_1 and p_2 denote the first two principal eigenvectors, then the initial weight of neuron j is defined by Equation 5 [30].

$$w_j^{(0)} = \bar{x} + a_j P_1 + b_j P_2 \quad (5)$$

a_i and b_j are the neuron position coefficients on the lattice. This initialization is used to reduce variability caused by random initialization and to accelerate early map organization.

2.5 DTW-based Representative Neuron Search

The main modification lies in the similarity measurement mechanism. In the proposed model, Euclidean distance in the BMU search phase is replaced by DTW. For two time series $x = \{x_1, x_2, \dots, x_T\}$ and $w = \{w_1, w_2, \dots, w_T\}$, DTW is defined as the minimum cost of the alignment path $P = \{(a_l, b_l)\}_{l-1}^{L}$ in Equation 6 [4].

$$DTW(x, w) = \min_P \sum_{l-1}^{L} d\left(x_{a_l}, w_{b_l}\right) \quad (6)$$

with local cost function $d\left(x_{a_l}, w_{b_l}\right) = \left(x_{a_l} - w_{b_l}\right)^2$. The computation is performed using the dynamic relation in Equation 7 [19].

$$D(i,j) = d\left(x_i, w_j\right) + min\{D(i-1,j), D(i,j-1), D(i-1,j-1)\} \quad (7)$$

The representative neuron is then determined on the basis of the minimum DTW value, as expressed in Equation 8 [4].

$$c = \arg\min_j \ DTW\left(x, w_j\right) \quad (8)$$

2.6 Multi-BMU Mechanism

The second main modification is the implementation of a multi-BMU mechanism. Unlike classical SOM, which selects only one BMU, the proposed model selects multiple BMUs, denoted by k. In this study, the three best neurons are selected as BMUs (k = 3) on the basis of their DTW ranking B_k (x)=$\{b_1,b_2,b_3\}$ dengan $DTW(x,w_{b1}) \leq DTW(x,w_{b2}) \leq DTW(x,w_{b3})$. The top-k strategy is used because it is more stable than a similarity-threshold approach, which is highly dependent on the distance distribution. The contribution of each winning neuron is then calculated using similarity-based weighting according to Equation 9 [28].

$$\omega_m = \frac{exp\left(-\beta\ DTW\left(x, w_{b_m}\right)\right)}{\sum_{q-1}^{3} exp\left(-\beta\ DTW\left(x, w_{b_q}\right)\right)} \quad (9)$$

$\beta>0$ is the sharpness parameter of the distribution. The smaller the DTW value, the greater the contribution of the neuron in the weight update process

2.7 Computational Complexity

The computational complexity of the proposed SOM–DTW–multi-BMU model is higher than that of the traditional SOM because it replaces Euclidean distance with DTW and extends winner selection from a single BMU to multiple BMUs. For the traditional SOM, the dominant cost of BMU search and weight updating is $O(N.M.L)$, leading to a total training complexity of $O(T.N.M.L)$ with T=50, N=2505, M=225 dan L=25, this become $O(50 \times 2505 \times 225 \times 25) = O(704531250) \approx O(7.05 \times 10^8)$.

In the improved SOM, PCA initialization adds a one-time cost of $O(NL^2)$, while DTW-based BMU search dominates the training cost with $O(N.M.L^2)$ The N-BMU selection adds $O(M \log M)$, and virtual-neuron formation contributes $O(k.L)$, both of which are relatively small compared with DTW. The total complexity is therefore dominated by $O(T.N.M.L^2)$ Using the same values, the estimated complexity becomes $O(50 \times 2505 \times 225 \times 625) = O(17613281250) \approx O(1.76 \times 10^{10})$ which is about 25 times larger than that of the traditional SOM.

To reduce this burden, the proposed model applies a DTW window constraint Sakoe–Chiba band, which restricts the warping path around the diagonal and

reduces the distance computation cost from $O(L^2)$ to approximately $O(Lr)$, where r is the window width and r≪L. Consequently, the dominant DTW search cost can be reduced from $O(NML^2)$ to approximately $O(NMLr)$. If a Sakoe–Chiba window width of r=5 is used, then the total computational complexity becomes $O(50 \times 2505 \times 225 \times 25 \times 5) \approx O(3.52 \times 10^9)$.

For clarity, the computational differences between the traditional SOM and the improved SOM are summarized in Table 1. The comparison shows that the main increase in computational cost comes from the replacement of Euclidean distance with DTW, whereas the multi-BMU mechanism contributes only a relatively small additional cost compared with the DTW search itself.

Table 1. Computational Complexity Comparison Between Traditional SOM and Improved SOM

Component	Traditional SOM	Improved SOM (Full DTW)	Improved SOM (DTW with Sakoe–Chiba band)
Initialization	(O(1))	(O(NL^2))	(O(NL^2))
Distance	(O(L))	(O(L^2))	(O(Lr))
Winner selection	1 BMU	(k)-BMU	(k)-BMU
Weight update	(O(NML))	(O(NML))	(O(NML))
Total complexity	(O(TNML))	(O(TNML^2))	(O(TNMLr))
Estimated operations	(7.05 x 10^8)	(1.76 x 10^{10})	(3.52 x 10^9) (for r=5)
Relative increase vs traditional SOM	–	(± 25x)	(± 5x)

As shown in Table 1, the application of a Sakoe–Chiba band substantially reduces the computational burden of the improved SOM. Although the constrained model remains more expensive than the traditional SOM, it is considerably more efficient than the full DTW-based version while still preserving the main benefit of alignment-aware temporal similarity.

3. Results and Discussions

3.1 U-Matrix Analysis

Training results were analyzed using U-Matrix maps to examine topological structure and visual separation among neuron regions. The U-Matrix of the conventional SOM and that of the improved model display noticeably different patterns, providing important visual context for interpreting the trade-off between clustering fidelity and topological fidelity.

Figure 2 visually shows that the conventional SOM produces a relatively broader distribution of inter-neuron distances and fails to form clear separation blocks among clusters. This is consistent with the low Silhouette score, indicating weak cohesion and separation among the resulting clusters. In contrast, Figure 3 shows that the modified model produces more structured distance gradients and clearer separation boundaries in several regions of the map. This pattern supports the finding that DTW strengthens sensitivity to temporal shape similarity [4], [22], while PCA initialization helps direct the initial organization of neurons more effectively [28], [30].

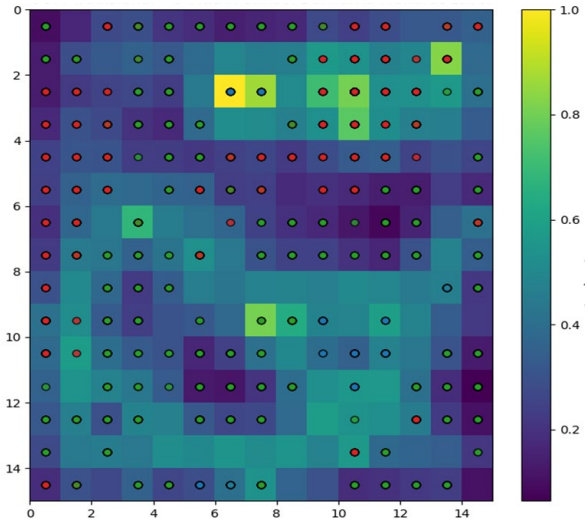


Figure 2. U-Matrix of the conventional SOM. The map shows a relatively more irregular distribution of neuron distances and a less consolidated spatial organization across the lattice

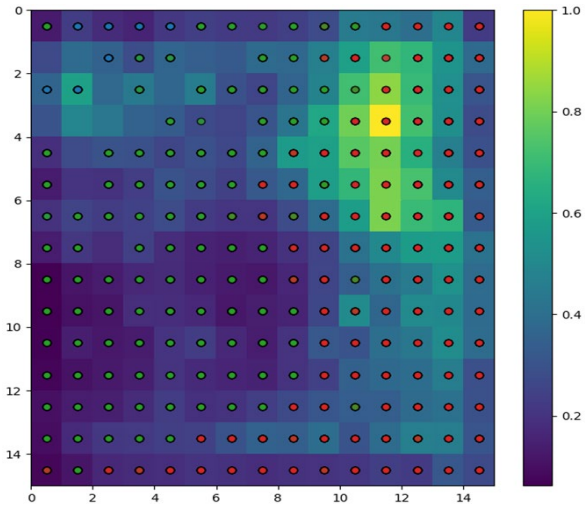


Figure 3. U-Matrix of the improved SOM with PCA-init and DTW-Multi-BMU. The map exhibits a more structured separation pattern, indicating stronger temporal grouping despite a less topology-conservative organization

However, the improved model also shows that map organization does not fully follow the smooth topological pattern characteristic of classical SOM. Several separation regions appear sharper and more discrete, which conceptually aligns with the increase in TE. Thus, the U-Matrix visualization reinforces the statistical interpretation that improvements in temporal cluster quality are achieved through a more discriminative but less topology-conservative map reorganization.

3.2 Comparative Performance Across Model

This section presents a comparative evaluation of four models: M1 (standard SOM), M2 (SOM + DTW), M3 (SOM + Multi-BMU), and M4 (hybrid SOM + DTW +

Multi-BMU). The evaluation was conducted across 10 independent runs using four metrics: the Silhouette score to assess clustering structure, TE to assess preservation of SOM topological structure, the Davies–Bouldin Index (DBI) evaluates clustering quality by measuring the average similarity between each cluster and the Calinski–Harabasz Index (CHI) measures the ratio of between-cluster dispersion to within-cluster dispersion. This comparative design is important because it makes it possible to isolate the contribution of each component, namely alignment-aware similarity through DTW and distributed winner adaptation through multi-BMU.

Table 2. Per-run Topographic Error for M1–M4

Run	M1: SOM standar	M2: SOM + DTW	M3: SOM + multi-BMU	M4: Hybrid SOM + DTW + multi-BMU
0	0.139721	0.150781	0.135589	0.156487
1	0.149301	0.163193	0.157405	0.172455
2	0.141317	0.169008	0.168993	0.204391
3	0.192415	0.179302	0.160599	0.187226
4	0.118164	0.165002	0.14798	0.171257
5	0.163273	0.163752	0.155363	0.164072
6	0.149301	0.181617	0.143562	0.186826
7	0.178044	0.181158	0.198607	0.183234
8	0.131737	0.196417	0.196815	0.211577
9	0.145309	0.166363	0.185948	0.200399

Table 2 presents the per-run TE values for all models. M1 generally produced the lowest TE values, whereas M4 produced the highest. This pattern indicates that the stronger the modifications introduced to improve temporal sensitivity and prototype-learning flexibility, the greater the potential reduction in topological preservation. The intermediate positions of M2 and M3 support a stepwise interpretation in which each additional component carries consequences for SOM topological structure.

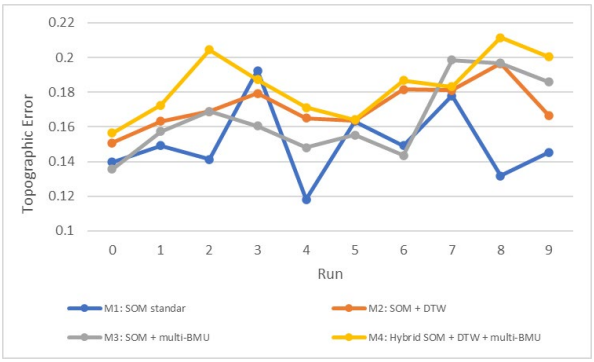


Figure 4. Comparison of TE across 10 runs for M1–M4

Figure 4 shows that M1 tends to produce the lowest TE values, whereas M4 tends to produce the highest. This indicates that when the model is made more sensitive to temporal similarity and more flexible in prototype updating, the classical topological behavior of SOM is no longer fully preserved. In particular, using DTW in M2 changes the geometry of the similarity space from Euclidean space to temporal alignment space.

As a result, the proximity relations learned by the model differ from those assumed by the conventional SOM lattice. Meanwhile, multi-BMU in M3 broadens the influence of winning neurons and makes prototype updating more distributed, which may reduce local topological rigidity. The fact that M3 performs better than M2 in TE is also noteworthy, suggesting that multi-BMU may be more topology-friendly than DTW when each is applied in isolation. However, when both are combined in M4, cluster quality improves the most, but at the cost of the highest TE among all models. Figure 5 further clarifies this mean-level pattern

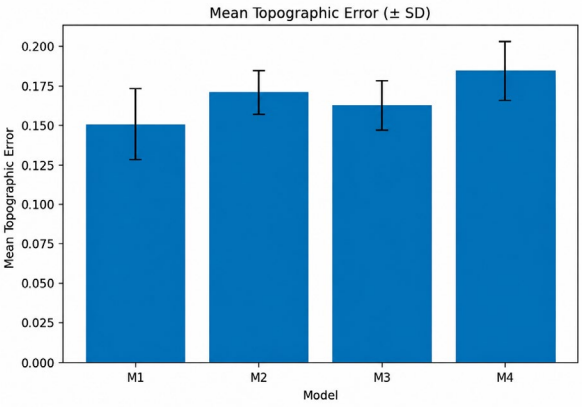


Figure 5. Mean TE (± SD) for M1–M4

Table 3. Per-run Silhouette Score for M1–M4

Run	M1: SOM standar	M2: SOM + DTW	M3: SOM + multi-BMU	M4: Hybrid SOM + DTW + multi-BMU
0	0.127703	0.623357	0.30867	0.752
1	0.127242	0.773261	0.38524	0.842
2	0.129534	0.548731	0.479163	0.746
3	0.125439	0.58279	0.521154	0.851
4	0.118971	0.483156	0.376622	0.797
5	0.127444	0.58622	0.34955	0.838
6	0.135565	0.605181	0.300085	0.76
7	0.112771	0.60275	0.328087	0.793
8	0.13271	0.402267	0.348001	0.54
9	0.121223	0.450311	0.300547	0.764

Table 3 shows a different pattern from Table 1: M4 consistently achieves the highest Silhouette score, whereas M1 produces the lowest. The intermediate positions of M2 and M3 indicate that both DTW and multi-BMU improve cluster quality relative to the baseline, but their combination in M4 yields the strongest performance. These results support the conclusion that the hybrid model is the most effective at generating temporal clusters that are both compact and well separated.

Figure 6 shows that adding DTW in M2 produces a larger gain in clustering quality than adding multi-BMU alone in M3. This suggests that, for time-series data, alignment-aware similarity modeling has a particularly strong influence on cluster quality. At the same time, M3 still performs clearly better than M1, indicating that distributed winner adaptation also adds value by improving how prototypes are formed in SOM. The strongest improvement appears in M4, which outperforms both M2 and M3. This shows that the best performance arises not from either component alone,

but from the synergy between more appropriate similarity measurement and richer prototype learning.

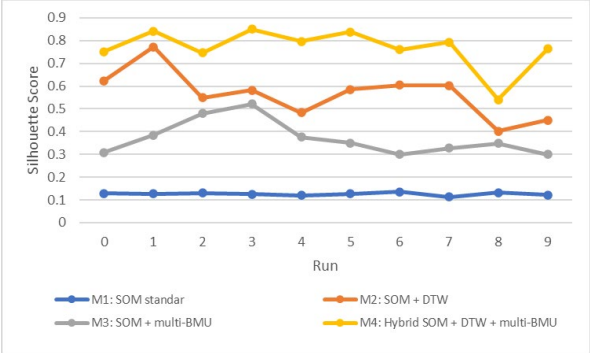


Figure 6. Comparison of Silhouette Score across 10 runs for M1–M4

To clarify the comparison of mean performance, Figure 7 presents a bar chart of mean ± standard deviation for the Silhouette score

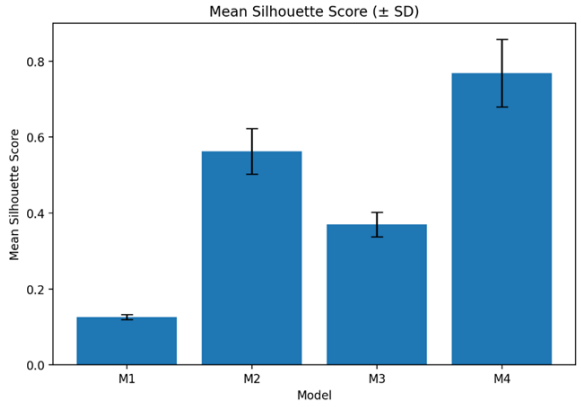


Figure 7. Mean Silhouette Score (± SD) for M1–M4

Figure 7 shows a very clear difference among the models. M1 has a very low mean value, whereas M4 substantially outperforms the others. From a methodological perspective, these results confirm that the main objective of this study—improving time-series clustering quality in SOM—is most clearly achieved by the hybrid model.

Table 4 presents the additional clustering validity results using the DBI and CHI. The results show a consistent pattern across the four models. The hybrid model M4 (SOM + DTW + multi-BMU) achieved the lowest DBI (0.4011) and the highest CHI (233.1750), indicating the best overall clustering quality. In contrast, the standard SOM M1 produced the highest DBI (4.6261) and the lowest CHI (38.2708), suggesting the weakest cluster compactness and separation.

Table 4. Clustering Validity Results using DBI and CHI

Model	Mean DBI	Mean CHI
M1	4.6261	38.2708
M2	2.5045	115.9852
M3	3.0021	198.234
M4	0.4011	233.1750

The intermediate results of M2 and M3 further clarify the contribution of each component. Based on DBI, M2 (SOM + DTW) outperformed M3 (SOM + multi-BMU), suggesting that DTW contributes more strongly to reducing cluster overlap. However, based on CHI, M3 outperformed M2, indicating that the multi-BMU mechanism contributes more strongly to improving global partition structure. Taken together, these findings suggest that DTW and multi-BMU improve clustering quality through different mechanisms, and their integration in M4 produces the strongest and most balanced effect.

3.3 Statistical Analysis

To ensure that performance differences among models were not merely descriptive, statistical analyses were conducted for TE and the Silhouette score. Because this study compares four models evaluated on the same runs, the analysis proceeded in two stages. First, the Friedman test was used as a global test to determine whether performance differed significantly across the four models. Second, when the Friedman test was significant, post-hoc pairwise comparisons were performed using the Wilcoxon signed-rank test. To control the risk of inflated Type I error due to multiple comparisons, the p-values from the Wilcoxon tests were adjusted using the Holm procedure.

This approach was selected because the data came from repeated paired runs and the number of observations was relatively small (n = 10), making non-parametric tests more appropriate than parametric alternatives. In interpreting the results, higher Silhouette scores indicate better cluster quality, whereas lower TE values indicate better topological preservation.

Table 5. Friedman test for overall comparison among M1–M4

Metric	Friedman χ^2	df	p-value	Interpretation
TE	16.56	3	0.00087	Significant difference among models
Silhouette Score	30	3	0.00000138	Significant difference among models

Based on Table 5, the Friedman test results show significant differences among the four models for both TE and the Silhouette score. Accordingly, the analysis was extended with post-hoc pairwise comparisons to identify which model pairs differed significantly.

Table 6. Pairwise Wilcoxon signed-rank test results for TE

Pair	Wilcoxon Statistic	Raw p-value	Holm-p-val	Significant ($\alpha = 0.05$)
M1 vs M2	4	0.01367	0.05469	No
M1 vs M3	14	0.19336	0.38672	No
M1 vs M4	2	0.00586	0.0293	Yes
M2 vs M3	18	0.375	0.38672	No
M2 vs M4	0	0.00195	0.01172	Yes
M3 vs M4	5	0.01953	0.05859	No

The results in Table 6 show that, after Holm correction, significant TE differences remain only for the pairs M1 vs M4 and M2 vs M4. This means that, in terms of topological preservation, the hybrid model M4 performs significantly worse than M1 and M2, but not significantly differently from M3.

Unlike TE, the results in Table 7 show that all model pairs differ significantly for the Silhouette score, even after Holm correction. This is a strong finding because it indicates that every modification added to SOM produces a real difference in clustering quality. In other words, DTW, multi-BMU, and their combination all exert statistically significant effects on cluster structure.

Table 7. Pairwise Wilcoxon signed-rank test results for Silhouette Score

Pair	Wilcoxon Statistic	Raw p-value	Holm-adjusted p-value	Significant ($\alpha = 0.05$)
M1 vs M2	0	0.00195	0.01172	Yes
M1 vs M3	0	0.00195	0.01172	Yes
M1 vs M4	0	0.00195	0.01172	Yes
M2 vs M3	0	0.00195	0.01172	Yes
M2 vs M4	0	0.00195	0.01172	Yes
M3 vs M4	0	0.00195	0.01172	Yes

3.4 Overall Interpretation of M1–M4

When interpreted step by step, M1 represents the baseline behavior of conventional SOM: the map preserves topological structure relatively well, but internal clustering quality remains low. The very low mean Silhouette score of M1 indicates that standard SOM is not sufficiently effective at producing compact and well-separated temporal clusters when the data exhibit asynchronous dynamics. This finding is consistent with the view that classical SOM, which relies on Euclidean distance and single-BMU learning, is more suitable for static vector data than for elastic temporal data [8], [19].

The transition from M1 to M2 shows a very clear change. When Euclidean distance is replaced by DTW, clustering quality improves sharply. This suggests that a major weakness of standard SOM in the time-series setting lies in how similarity is defined. In temporal data, two substantively identical patterns may appear with phase shifts, local distortions, or different rates of evolution. Because Euclidean distance compares points directly, it often fails to recognize this type of similarity. DTW addresses the problem by allowing nonlinear alignment between sequences, so that shape similarity can still be captured even when timestamps are not perfectly aligned. This finding is highly consistent with [19], which showed that DTW-SOM is more effective than Euclidean SOM in identifying patterns in environmental exposure data.

The transition from M1 to M3 shows a different pattern. Adding multi-BMU also improves clustering quality, but the gain is not as large as that obtained by adding DTW. This suggests that extending the winning mechanism from one neuron to several indeed enriches prototype learning, but its main strength lies not in similarity definition but in how prototypes are formed and adapted. In temporal data, this is reasonable because a single time series often cannot be adequately represented by one winning neuron alone. Allowing several nearby neurons to contribute to updating provides more flexible representations of local temporal variation. This finding supports the argument in [28] that virtual winning neurons or multi-winner mechanisms can improve SOM learning by reducing dependence on a single winner.

The position of M4 as the best-performing model in terms of the Silhouette score indicates that the greatest performance improvement does not come from either component alone, but from the synergy between DTW and multi-BMU. In other words, for SOM applied to temporal data, similarity measurement and prototype learning are closely interdependent. DTW helps the model identify neurons that are genuinely similar in temporal terms, whereas multi-BMU enables prototypes to adapt more richly to complex input structure. The superiority of M4 therefore reflects not a simple sum of the effects of M2 and M3, but the interaction of two mutually reinforcing mechanisms

3.5 Limitations and Future Work

Although the results of this study are robust, several limitations should be acknowledged. First, the experiment was conducted in only one application domain, namely room-lighting data, so cross-domain generalization still needs to be examined. Second, although the number of runs is sufficient for non-parametric inferential analysis, using more datasets would strengthen the generalizability of the findings. Third, while the hybrid model clearly improves cluster quality, it does not explicitly resolve the increase in TE, leaving room for further methodological refinement. The study therefore contributes not only by proposing a stronger model, but also by clarifying the boundaries and consequences of that development. In other words, the superiority of M4 is built on a shift in SOM orientation from topology-centered mapping toward a temporally sensitive clustering framework, and this insight contributes more meaningfully to the SOM literature for temporal data than a mere report of higher scores [23], [32].

Several directions for future work may be pursued. First, the model should be tested on more diverse temporal datasets in terms of domain, sequence length, and pattern complexity to evaluate generalizability more rigorously. Second, future studies may explore different values of k in the multi-BMU mechanism, including the possibility of adaptive-k strategies in which the number of winners is dynamically adjusted to input characteristics. Third, methodological efforts are needed to reduce TE without sacrificing clustering quality, for example through adaptive neighborhood functions, topological regularization, or staged updating schemes. Fourth, because DTW entails higher computational cost, future work should consider lower bounding, pruning, or warping-window optimization to improve scalability on larger datasets. Overall, this study suggests that the hybrid SOM-DTW-Multi-BMU framework is a promising direction for time-series clustering, particularly when the primary analytical goal is to obtain temporal clusters that are more compact, better separated, and more substantively meaningful.

4. Conclusions

This study proposed a hybrid SOM–DTW–multi-BMU framework for time-series clustering of indoor lighting data. The results show that the hybrid model (M4) achieved the best clustering performance, with the highest mean Silhouette score (0.768300), the lowest DBI (0.4011), and the highest CHI (233.1750). Compared with the baseline M1, whose mean Silhouette score was 0.125860, M4 improved clustering quality by 0.642440 points, or approximately 510.44%. However, M1 still produced the lowest mean TE (0.150858), indicating better topology preservation.

These findings indicate that DTW and multi-BMU improve SOM through different but complementary mechanisms. DTW strengthens temporal similarity modeling, whereas multi-BMU enriches prototype adaptation. Their integration in M4 produces the strongest overall effect on cluster compactness and separation. From a computational perspective, this improvement is obtained at the cost of higher complexity, as the model shifts from the linear Euclidean-based SOM setting toward a DTW-based temporal matching scheme. Nevertheless, the added cost remains methodologically justified because it yields substantially better temporal clustering quality, and the use of a DTW window constraint helps control the computational burden. Therefore, the main contribution of this study lies in showing that alignment-aware similarity and distributed winner adaptation should be jointly designed within a single SOM framework for more effective temporal clustering.

Declarations

AI Usage Statement

The authors used chatGPT only for language refinement and improvement of manuscript readability; all scientific content, analysis, and conclusions are the authors' own. The authors also used deepL for translate from Bahasa to English.

Author Contributions

Rina Dewi Indahsari: conceptualization, methodology, model development, implementation, formal analysis, visualization, writing – original draft, and writing – review and editing.

Siti Sendari: supervision, methodology validation, writing – review and editing, and project administration.

Wahyu Sakti Gunawan Irianto: supervision, methodology validation, computational analysis, and writing – review and editing.

Samsul Arifin: IoT development, data collecting, data analysis, and technical validation.

Diah Lestari: data curation, visualization support, and writing – review and editing.

Funding Statement

This research received no external funding.

Conflict of Interest Statement

The authors declare no conflict of interest regarding the publication of this research.

Ethical Clearance

This study does not involve human participants, animal subjects, or interventions requiring ethical approval. The research was conducted using computational experiments on indoor illuminance time-series data collected from an office building environment. All procedures related to data processing and analysis were performed in accordance with applicable academic and research integrity standards.

Data and Code Availability

The dataset and source code used in this study are not publicly available because they are derived from specific indoor environmental measurements and experimental implementations. However, the materials may be made available from the corresponding author upon reasonable request for academic and research purposes.

References

[1] A. Alqahtani, M. Ali, X. Xie, and M. W. Jones, "Deep Time-Series Clustering: A Review," *Electronics*, vol. 10, no. 23, p. 3001, Dec. 2021, doi: 10.3390/electronics10233001.

[2] C. Ji *et al.*, "Time series classification with random temporal features," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 9, p. 101783, Oct. 2023, doi: 10.1016/j.jksuci.2023.101783.

[3] Ł. Borchmann, D. Jurkiewicz, F. Graliński, and T. Górecki, "Dynamic Boundary Time Warping for sub-sequence matching with few examples," *Expert Systems with Applications*, vol. 169, p. 114344, May 2021, doi: 10.1016/j.eswa.2020.114344.

[4] A. Javed, D. M. Rizzo, B. S. Lee, and R. Gramling, "Somtimes: self organizing maps for time series clustering and its application to serious illness conversations," *Data Mining and Knowledge Discovery*, vol. 38, no. 3, pp. 813–839, 2024, doi: 10.1007/s10618-023-00979-9.

[5] A. Wibawa, A. B. Putra Utama, H. Elmunsyah, U. Pujianto, F. Dwiyanto, and L. Hernandez, "Time-series analysis with smoothed Convolutional Neural Network," *Journal of Big Data*, vol. 9, p. 44, Apr. 2022, doi: 10.1186/s40537-022-00599-y.

[6] Q.-V. Doan, H. Kusaka, T. Sato, and F. Chen, "S-SOM v1.0: a structural self-organizing map algorithm for weather typing," 2021, Accessed: Sep. 18, 2025. doi: 10.5194/gmd-14-2097-2021

[7] A. Pranolo *et al.*, "Enhanced Multivariate Time Series Analysis Using LSTM: A Comparative Study of Min-Max and Z-Score Normalization Techniques," *ILKOM Jurnal Ilmiah*, vol. 16, no. 2, pp. 210–220, Aug. 2024, doi: 10.33096/ilkom.v16i2.2333.210-220.

[8] J. K. Krishnan and K. Mitra, "A modified Kohonen map algorithm for clustering time series data," *Expert Systems with Applications*, vol. 201, p. 117249, Sep. 2022, doi: 10.1016/j.eswa.2022.117249.

[9] A. W. Saputra, A. P. Wibawa, U. Pujianto, A. B. P. Utama, and A. Nafalski, "LSTM-based Multivariate Time-Series Analysis: A Case of Journal Visitors Forecasting," *ILKOM Jurnal Ilmiah*, vol. 14, no. 1, pp. 57–62, Apr. 2022, doi: 10.33096/ilkom.v14i1.1106.57-62.

[10] L. E. Ekemeyong Awong and T. Zielinska, "Comparative Analysis of the Clustering Quality in Self-Organizing Maps for Human Posture Classification," 2023, Accessed: May 06, 2026. doi: 10.3390/s23187925
[11] P. H. Franses and T. Wiemann, "Intertemporal Similarity of Economic Time Series: An Application of Dynamic Time Warping," *Computational Economics*, vol. 56, no. 1, pp. 59–75, Apr. 2020, doi: 10.1007/s10614-020-09986-0.
[12] S. Sendari, A. B. P. Utama, N. S. F. Putri, P. Widiharso, and R. J. Putra, "K-Means and Fuzzy C-Means Optimization using Genetic Algorithm for Clustering Questions," *International Journal of Advanced Science and Computer Applications*, vol. 1, no. 1, pp. 1–9, Jan. 2022, doi: 10.47679/ijasca.v1i1.2.
[13] A. P. Wibawa *et al.*, "Mean-Median Smoothing Backpropagation Neural Network to Forecast Unique Visitors Time Series of Electronic Journal," *Journal of Applied Data Sciences*, vol. 4, no. 3, pp. 163–174, Sep. 2023, doi: 10.47738/jads.v4i3.97.
[14] L. Bermúdez, I. Morillo, and A. Salazar, "Supervised clustering using SOM for severity-based pattern detection in urban traffic crashes," *Expert Systems with Applications*, vol. 316, p. 131895, Jun. 2026, doi: 10.1016/j.eswa.2026.131895.
[15] X. Li *et al.*, "Integration of self-organizing map and Monte Carlo simulation for ecological risk prediction of heavy metal attenuation in groundwater," *Ecotoxicology and Environmental Safety*, vol. 302, p. 118761, Sep. 2025, doi: 10.1016/j.ecoenv.2025.118761.
[16] K. Li *et al.*, "Using dynamic time warping self-organizing maps to characterize diurnal patterns in environmental exposures," *Scientific Reports*, vol. 11, no. 1, p. 24052, Dec. 2021, doi: 10.1038/s41598-021-03515-1.
[17] M. C. G. Kopland and E. J. Giltay, "Dynamic Time Warp (DTW) as a scalable, data-efficient, and clinically relevant analysis of dynamic processes in patients with psychiatric disorders: a tutorial," *Journal of Eating Disorders*, vol. 13, no. 1, p. 230, Oct. 2025, doi: 10.1186/s40337-025-01414-8.
[18] S. Sendari *et al.*, "Common-Sensical Incentive Reward in Deep Actor-Critic Reinforcement Learning for Mobile Robot Navigation," 2024, *ICIC International 学会*: 02. doi: 10.24507/ijicic.20.02.373.
[19] K. Li *et al.*, "Using dynamic time warping self-organizing maps to characterize diurnal patterns in environmental exposures," *Scientific Reports*, vol. 11, no. 1, p. 24052, Dec. 2021, doi: 10.1038/s41598-021-03515-1.
[20] L. Qiu, C. Qiu, and C. Song, "ESDTW: Extrema-based shape dynamic time warping," *Expert Systems with Applications*, vol. 239, p. 122432, Apr. 2024, doi: 10.1016/j.eswa.2023.122432.
[21] N. Kumar, S. K. Sood, and M. Saini, "Internet of Vehicles (IoV) based Framework for Vehicle Degradation using Multidimensional Dynamic Time Warping (MDTW)," *Expert Systems with Applications*, vol. 224, p. 120038, Aug. 2023, doi: 10.1016/j.eswa.2023.120038.
[22] N. Vaughan, "Nearest centroid classifier for multivariate time series using DTW-MP centroid," *Expert Systems with Applications*, vol. 318, p. 132069, Jul. 2026, doi: 10.1016/j.eswa.2026.132069.
[23] N. P. Rougier and G. Is. Detorakis, "Randomized Self-Organizing Map," *Neural Comput*, vol. 33, no. 8, pp. 2241–2273, Jul. 2021, doi: 10.1162/neco_a_01406.
[24] T. Widiyaningtyas, I. A. E. Zaeni, and P. Y. Wahyuningrum, "Self-Organizing Map (SOM) For Diagnosis Coronary Heart Disease," *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, pp. 286–289, Nov. 2019, doi: 10.1109/ICITISEE48480.2019.9003746.
[25] G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, M. Re, and S. Spano, "AW-SOM, an Algorithm for High-Speed Learning in Hardware Self-Organizing Maps," *IEEE Trans. Circuits Syst. II*, vol. 67, no. 2, pp. 380–384, Feb. 2020, doi: 10.1109/TCSII.2019.2909117.
[26] H. Mayatopani and N. Handayani, "Implementation of Self-Organizing Map (SOM) Algorithm for Image Classification of Medicinal Weeds," Dec. 2023, Accessed: Apr. 13, 2024. doi: 10.29207/resti.v7i3.4755
[27] N. M. Hai, T. V. Lang, and T. The Van, "Improving the Efficiency of Semantic Image Retrieval Using a Combined Graph and SOM Model," *IEEE Access*, vol. 11, pp. 140646–140659, 2023, doi: 10.1109/ACCESS.2023.3333678.
[28] X. Fan, S. Zhang, X. Xue, R. Jiang, S. Fan, and H. Kou, "An Improved Self-Organizing Map (SOM) Based on Virtual Winning Neurons," *Symmetry*, vol. 17, no. 3, p. 449, Mar. 2025, doi: 10.3390/sym17030449.
[29] S. Jamil *et al.*, "Multi Self-Organizing Map (SOM) Pipeline Architecture for Multi-View Clustering," *IEEE Access*, vol. 12, pp. 85806–85821, 2024, doi: 10.1109/ACCESS.2024.3412950.
[30] A. Abdelaziz, V. Santos, M. S. Dias, and A. N. Mahmoud, "A hybrid model of self-organizing map and deep learning with genetic algorithm for managing energy consumption in public buildings," *Journal of Cleaner Production*, vol. 434, p. 140040, Jan. 2024, doi: 10.1016/j.jclepro.2023.140040.
[31] J. Deng *et al.*, "Potential Global Invasion Risk of Scale Insect Pests Based on a Self-Organizing Map," *Insects*, vol. 14, no. 7, p. 572, Jun. 2023, doi: 10.3390/insects14070572.
[32] S. Licen, A. Astel, and S. Tsakovski, "Self-organizing map algorithm for assessing spatial and temporal patterns of pollutants in environmental compartments: A review," *Science of The Total Environment*, vol. 878, p. 163084, Jun. 2023, doi: 10.1016/j.scitotenv.2023.163084.