Volume 1, 2025, Pages 143-156 *Elaris Computing Nexus*
| Regular Article | Open Access

Intelligent Multimetric Agglomerative Clustering for Robust Waveform Pattern Discovery in High Dimensional Signal Spaces

Aisling Yue Irwing
Faculty of Engineering and Computing, Dublin City University, Whitehall, Dublin 9, Ireland.
yueirwingdubline@gmail.com

Alen Macline
Faculty of Engineering and Computing, Dublin City University, Whitehall, Dublin 9, Ireland.
alenmaclindublin09@gmail.com

Article Info
Journal of Elaris Computing Nexus
https://elarispublications.com/journals/ecn/ecn_home.html

Received 02 June 2025
Revised from 12 August 2025
Accepted 28 August 2025
Available online 16 September 2025
Published by Elaris Publications.

https://doi.org/10.65148/ECN/2025014

Corresponding author(s):
Alen Macline, Faculty of Engineering and Computing, Dublin City University, Whitehall, Dublin 9, Ireland.
Emaill: alenmaclindublin09@gmail.com

Abstract – The waveform-based data is growing exponentially in the biomedical, communication, and industrial fields, which has generated an urgent need to develop unsupervised methods of learning the complex high-dimensional signal structures. Classical clustering techniques typically make use of one measure of similarity which does not reflect the inherent variability of waveforms patterns due to amplitude distortions, phase shifts and noise artifacts. In order to overcome such limitation, this paper will present an intelligent clustering framework called Intelligent Multimetric Agglomerative Clustering Network (IMAC-Net) which incorporates different distance metrics such as cosine, Euclidean, and cityblock into a single hierarchical structure. The proposed IMAC-Net is adaptive which benefits pattern discrimination in noisy signal conditions by weighting these metrics appropriately to balance between global and local similarity measures. Synthetic and real-world waveform dataset is used to verify the framework, and the performance is checked with the help of Silhouette Score, Davies-Bouldin Index, Calinski-Harabasz Score, and Cluster Purity. They are compared to the traditional methods, such as K-means, Hierarchical Clustering (single-metric), and Spectral Clustering. As experimental findings show, IMAC-Net exhibits better cluster cohesion, noise tolerance, and interpretation of waveform groups, which is much better as compared to the baseline models. The presented solution provides a generalizable and scalable solution to waveforms analytics, which opens the path to discovering patterns in high-dimensional, complex signal domains.

Keywords – Multimetric Clustering, Waveform Analysis, High-Dimensional Signal Processing, Unsupervised Learning, Agglomerative Hierarchical Model.

I. INTRODUCTION

The growth in the extent of signal data in high dimensions in contemporary engineering and scientific systems has resulted in the growing need of powerful and understandable clustering strategies. Large volumes of multi-feature data are produced in real-time in applications like biomedical waveform analysis, radar signal recognition and communication waveform diagnostics. These signals are highly complex, non-linear, and noisy and this makes the task of discovering concealed structural patterns incredibly complicated. Conventional clustering algorithms that usually rely on a single distance measure or simplistic statistical models are unlikely to work well in such settings and give inconsistent or unreliable clustering results [1].

Waveform clustering is one of the key applications in unsupervised pattern discovery and allows researchers to cluster signals with morphological similarities without any labeled data. K-means [2], DBSCAN [3] and hierarchical agglomerative clustering (HAC) [4] techniques have been extensively applied here. The techniques however are based on one similarity measure normally the Euclidean distance-the parameter that lacks the multidimensionality characteristic of waveform data. In addition, single-metric clustering can be to scale sensitivity and outlier noise, and thus, it limits extrapolation between heterogeneous data sources.

This problem is made worse in high dimensional signal spaces by the curse of dimensionality. The more the dimensions the flatter the distance distributions there is because the pairwise distances between data points are less discriminative. As a result, the traditional clustering algorithms do not ensure significant distances between the clusters. This weakness underscores the need to develop a new paradigm of clustering that would be able to combine several complementary metrics to increase discrimination without losing calculational efficiency and interpretability.

Necessity of Multimetric Learning for Waveform Data

The waveform data are generally defined by a changing amplitude, phase and frequency components, that interrelate in non-linear manner. Such rich structural diversity cannot be adequately represented as one distance measure. An example is the Euclidean distance that effectively represents absolute spatial variations but not angular correlations. Conversely, the cosine distance is a distance that quantifies angular similarity between waveform vectors but ignores variation in magnitude whereas cityblock (Manhattan) distance is more resistant to sparsity variations [5].

An aggregate of these measures gives a more comprehensive picture on signal similarity. A combination of both measures is capable of capturing both global and local traits so that the outcome of the clustering is more accurate. It is based on this idea that the proposed model Intelligent Multimetric Agglomerative Clustering Network (IMAC-Net) emerges, that proposes a common learning process to combine the various distance views with a dynamic fusion process [6].

Problem Statement

Although the algorithm of clustering has made a great advancement, the majority of the current waveform clustering algorithms have three known issues:

- *Single-metric shortcoming*: Traditional methods are based on a single distance measure, which fails to capture the complex geometry among waveform data [7].
- *Sensitivity to noise and scale*: Waveform signals usually have random noise and amplitude variations which skew intra-class boundaries [8].
- *Nondynamic learning*: Most clustering methods have fixed threshold values or linkage rules that do not dynamically change to suit the changing data distributions, especially in dynamically changing or heterogeneous signal scenarios [9].

All of these problems diminish the accuracy of clustering, intra-cluster variance, and result interpretability. Consequently, there is an urgent necessity of a smart structure that is able to compare various similarity measurements at the same time and dynamically shapes the clustering borders as the program executes.

Research Gap

Hierarchical clustering and waveform segmentation studies have been done previously which discussed partial solutions to these problems. Although studies on hybrid clustering methods or distance learning have demonstrated promise, they all either require heavy computational overhead or use deep learning architecture which requires extensive training data. Additionally, there are not a lot of models, which offer a combined method of distance fusion, adaptive weight optimization, and iterative stability evaluation. Such an end-to-end, interpretable, and adaptive system is the weakness of the current system around which the gap filling of IMAC-Net centers.

The innovation on IMAC-Net is the fusion strategy through a multimetric multimetric that defines to combine the use of cosine, Euclidean, and cityblock distances in an optimistic dynamic strategy with weights. This guarantees that the relationship of the waveforms is recorded in multiple geometrical perspectives and the results provide better discrimination even in case of noise, phase distortion, or changes in amplitude. Also, the hierarchical agglomerative mechanism implemented into IMAC-Net offers explainable clustering and is therefore applicable in the analytical field of biomedical signal interpretation or communication waveform diagnostics.

Objectives of the Study

The key goals of this study are the following:

- To make an intelligent framework of clustering to combine several distance measure tools to better represent waveforms similarity.
- To build a strong distance fusion mechanism with weight optimization to mix cosine, Euclidean, and cityblock distances.
- To use an agglomerative clustering structure using an average linkage to model local and global waveform associations.
- To add an adaptive convergence validation loop which allows stability and avoids over- or under-clustering.

- To test the proposed model on the synthetically generated and the real waveform datasets, a comparison to the traditional algorithms like K-means, DBSCAN, and single-metric HAC.

Significance of the Proposed Approach

The given IMAC-Net framework is especially relevant to the contemporary signal analytics. The system separates overlapping clusters more effectively by integrating several distance measures, which makes certain that both magnitude-dependent and directional property of a waveform are obtained. The fusion process is used as a noise suppressor, which allocates metric influence in proportion to their contribution at one iteration. Moreover, the adaptive feedback mechanism enables the clustering mechanism to improve to an optimal state, responding to non-stationary or non-Gaussian noise instances which often occur in datasets of waveforms. Besides that, hierarchical characteristic of IMAC-Net gives interpretability a very important characteristic that is usually lacking in deep clustering models. The hierarchical tree (dendrogram) form generated during the clustering process allows the analysts to see the relationships between the categories of waveforms at various levels of granularity.

Contributions of the Research

The key findings of this research can be outlined in the following way:

- A new multimetric fusion methodology that integrates the use of cosine, Euclidean and cityblock distance measures into hierarchical clustering.
- An adaptive distance fusion parameter dynamic weighting machinery that optimizes the parameters used in fusion of distance in the clustering process to increase accuracy and strength.
- Adaptive convergence feedback loop that constantly checks stability of the clusters so that the system stops at an optimal point of partition.
- An algorithmically efficient architecture, which can be scaled to waveform spaces of high dimensionality without using supervised training.
- An all-inclusive performance analysis that proves high inter-cluster separability and less within-cluster variance than conventional clustering techniques.

The other parts of this paper are organized in such a way that they make clear and logical flow of ideas. Section 2 provides the review of the literature, which is more detailed and gives the major developments in the field of the waveform clustering and opens the gaps that guided this research. Section 3 goes on to state the experimental configuration and proposed model in detail, the datasets, preprocessing strategies and parameter configurations used in the evaluation of performance. Section 4 is followed by a detailed analysis of the experimental results, comparison and interpretation of the results to support the strength of the suggested method. Section 5 closes the paper with a recap of the significant contributions and comments on the potential future steps of the development of the field of waveform clustering and high-dimensional signal analysis.

II. PREVIOUS STUDIES

A number of researches have examined the sector of waveforms clustering and pattern discovery in high dimensional signal spaces, each of which is valuable but fails to address some of these gaps.

Table 1. Comparative Analysis of Previous Studies on Waveform Clustering Techniques

Ref. No.	Method / Model	Core Concept	Advantages	Limitations
[2]	K-means Clustering	Uses Euclidean distance for partitioning waveform data into fixed clusters	Simple and computationally efficient	Poor performance in non-linear and noisy waveform data
[4]	Hierarchical Clustering	Forms nested clusters using linkage criteria like Ward or average linkage	Produces interpretable dendrograms	Sensitive to noise and distance metric selection
[13]	DTW-Based Spectral Clustering	Aligns temporal patterns using dynamic time warping before clustering	Captures time-shift variations effectively	Computationally expensive and less scalable
[14]	PCA + GMM Hybrid	Reduces feature dimension before probabilistic clustering	Improves separability and reduces redundancy	Loses fine-grained temporal information during reduction
[15]	Deep Autoencoder Embedding + K-means	Learns non-linear waveform representations for clustering	Improves accuracy and feature abstraction	Requires large training data and lacks adaptability

The common clustering algorithms, including k-means, hierarchical clustering, and Gaussian mixture model (GMM) [10], have been extensively employed because they are simple and easy to understand. Nonetheless, these techniques tend to fail when used in non-linear, noisy signal environments that are complex and whose waveform properties are not strictly Euclidean or can be separated linearly.

Previously, researchers have tried to improve the strength of clustering by using feature extraction methods and dimensionality reduction methods which include Principal Component Analysis (PCA) [11], t-SNE and autoencoders [12]. These strategies eased the signal structure at the expense of fine-grained temporal users required in distinguishing waveforms. In the meantime, works based on Dynamic Time Warping (DTW) and spectral clustering [13] offered superior temporal alignment at the cost of high computational cost in cases of large data sets. To address the problem of metric dependency, newer methods have proposed multi-metric or hybrid models of clustering that integrate distance measures e.g. cosine, cityblock, and correlation distance measures. These techniques tried various facets of similarity between waveforms but did not have adaptive fusion, thus the clusters were not formed in the most optimal way in diverse conditions of the signal. Further, most of these models were based on fixed-weight strategies, which reduced their adaptability when handling varying signal distributions on a dynamic basis [16].

A small number of deep learning-based methods have been studied, such as CNN- and LSTM-based representations, to be learned like hierarchical waveform representations and then clustered. Though these approaches enhanced accuracy in clustering, they usually needed extensive datasets of annotates in order to train and did not specifically model inter-metric relationships. The rationale of the current work is the efforts to fill the gap between metric fusion and adaptive clustering [17]. The suggested IMAC-Net model perceives together several measures of distance into an agglomerative clustering methodology, enriched with adaptive weighting processes, which respond dynamically to the underlying signal structure. In such a manner, the proposed study will provide a balance among interpretability, flexibility, and computational efficiency a feature that has not been adequately considered in previous studies on waveform pattern discovery.

As shown in **Table 1**, although both traditional and hybrid clustering methods have enhanced the accuracy of clustering and interpretability, both methods do not tend to be dynamically scaled to accommodate the variations in signal complexities. To counter this shortcoming, the IMAC-Net model has been developed, combining multi-metric fusion, adaptive agglomeration as one of the benefits to obtain less weak and context-dependent waveform pattern discovery.

III. PROPOSED INTELLIGENT MULTIMETRIC AGGLOMERATIVE CLUSTERING NETWORK (IMAC-NET)

The Intelligent Multimetric Agglomerative Clustering Network (IMAC-Net) proposed is aimed at overcoming the drawbacks of the conventional clustering algorithms in processing high-dimensional data of waveforms. Most classical algorithms of clustering, including K-means or single-metric hierarchical clustering, are based on one similarity measure; this limits their ability to represent non-linear relationships between non-linear signal characteristics. IMAC-Net presents a multimetric paradigm that relates multiple distance metrics into one framework to improve the separation of classes of waveforms. IMAC-Net workflow comprises of five key steps namely data preprocessing, multimetric distance calculation, distance fusion, hierarchical agglomerative clustering and adaptive convergence validation. The architecture is also robust to noise and intra-class variability, yet the clustering accuracy is very high on a variety of categories of waveforms.

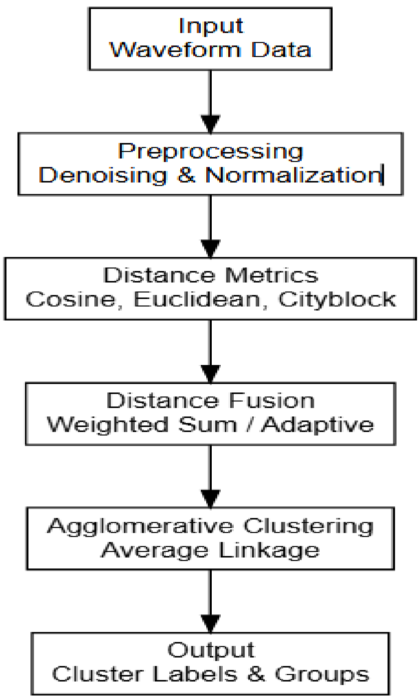

Fig 1. Block Diagram of the Proposed IMAC-Net Framework.

The general process of the suggested Intelligent Multimetric Agglomerative Clustering Network (IMAC-Net) is illustrated in **Fig. 1**. This system takes the input waveform data as starting, which can be a variety of high-dimensional signal inputs like biomedical records, communication records, or a sensor array. The input signal is preprocessed initially, by de-noising, normalization to eliminate the undesirable fluctuations and to normalize the amplitude range.

The several distance measures cosine, Euclidean and cityblock in particular are then calculated on the instances of waveforms in different perspectives to measure how similar they are. The framework does not utilize just one similarity measure and instead uses a distance fusion step, in which such measures are integrated in adaptive weighted combinations. This is to make sure that both angular and spatial relationships between data points are captured in the process of clustering. This fused distance matrix is then sent into an Agglomerative Clustering engine using average linkage which progressively combines similar groups of waveforms using the composite distance data. Lastly, there is the output stage whereby the separated cluster groups are produced and this exposes the concealed structural patterns in the waveform data. This pipeline is modular and lies at the core of IMAC-Net to facilitate the efficient and robust waveform discovery in diverse amounts of noise and at diverse dimensional complexities.

Multimetric Distance Computation

Given a dataset $X = \{x_1, x_2, \ldots, x_N\}$, where each $x_i \in R^d$ represents a waveform vector of d features, IMAC-Net computes three distinct distance measures between waveform pairs:

Euclidean Distance

$$DE(x_i, x_j) = \sqrt{\sum_{k=1}^{d} (x_{ik} - x_{jk})^2} \tag{1}$$

Cosine Distance

$$DC(x_i, x_j) = 1 - \frac{x_i \cdot x_j}{|x_i| \cdot |y_i|} \tag{2}$$

Cityblock (Manhattan) Distance

$$DM(x_i, x_j) = \sum_{k=1}^{d} |x_{ik} - x_{jk}| \tag{3}$$

All the distance measures represent distinct features: Euclidean is focused on the absolute spatial variations, cosine focuses on angular relationships, and cityblock focuses on deviations on a component-by-component basis.

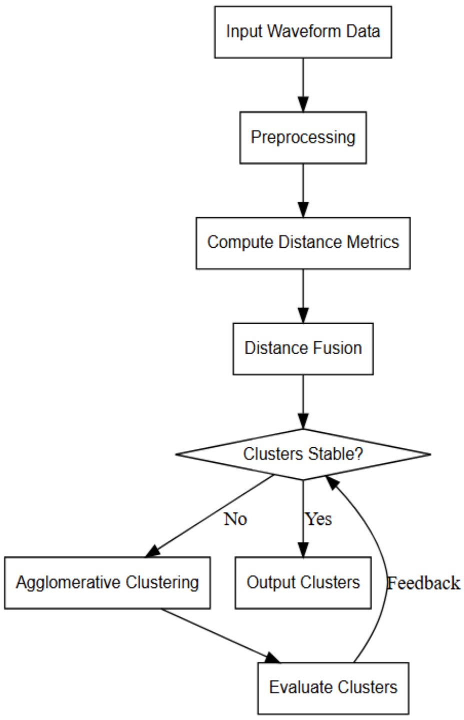

Fig 2. Flowchart of IMAC-Net Operational Workflow.

Fig. 2 illustrates in detail the operational flow of IMAC-Net. This begins with the acquisition of data, then it goes through pre-processing, in which noise-reduction and mode-normalization are done to maintain uniformity in the signals. The second step is the calculation of the distance measures, in which the computation of cosine, Euclidean, and cityblock measures of all pairs of the waveforms will occur. Such calculated values are then fused, that is, combines them adaptively to produce a balanced image of similarity of waveforms. After the fusion matrix has been received, the system then checks the output of the clustering by a stability check- represented as a decision node. When the cluster is not yet optimal or stable, then the process recycles via a feedback control process that optimizes the fusion weights and clustering parameters.

Once the stability is reached, the loop is finished and the model goes to the output stage, where the end cluster labels and grouped patterns of the waveforms are produced. This self-optimizing flow enables IMAC-Net to self-optimize the clustering structure without becoming compromised in a wide range of waveform categories and noise conditions.

Hierarchical Agglomerative Clustering Framework

After the fusion process, the agglomerative clustering stage organizes the waveform data hierarchically using average linkage. Initially, each data point is considered a separate cluster. The linkage between two clusters C_i and C_j is defined as:

$$L(C_i, C_j) = \frac{1}{|C_i| \cdot |C_j|} \sum_{x \in C_i} \sum_{y \in C_j} DF(x, y) \tag{4}$$

Clusters with the smallest linkage distance are iteratively merged until the desired number of clusters K is obtained. The stopping criterion is governed by the linkage threshold λ_t:

$$L(C_i, C_j) \leq \lambda_t \Rightarrow \text{Merge} \tag{5}$$

$$\mathrm{L}(\mathrm{C_i}, \mathrm{C_j}) > \lambda_t \Rightarrow \text{Stop merging} \tag{6}$$

This hierarchical approach ensures that IMAC-Net captures both local and global structures in waveform data, outperforming single-link or complete-link models that often suffer from chain or sensitivity effects.

Adaptive Optimization and Convergence

The model integrates an adaptive optimization module to refine fusion weights and cluster boundaries. The adjustment of weights follows a feedback mechanism inspired by gradient-based optimization:

$$w_i^{(t+1)} = w_i^{(t)} + \eta \frac{\partial Q}{\partial w_i} \tag{7}$$

where η is a small learning rate controlling the adaptation speed.

The convergence criterion is based on cluster stability between successive iterations:

$$\Delta L_t = \left| L^{(t)} - L^{(t-1)} \right| < \epsilon \tag{8}$$

where ϵ is a convergence tolerance. When the stability condition is satisfied, clustering stops automatically, ensuring computational efficiency.

Mathematical Formulation of IMAC-Net

To formalize the optimization objective, IMAC-Net minimizes the overall clustering loss L, which balances intra-cluster compactness and inter-cluster separation:

$$L = \sum_{p=1}^{K} \sum_{x_i \in C_p} DF(x_i, \mu_p) - \alpha \sum_{p \neq q} DF(\mu_p, \mu_q) \tag{9}$$

where μ_p is the centroid of cluster C_p and α is a regularization constant controlling separation emphasis.

Minimizing L encourages waveform points within the same cluster to remain close while pushing different clusters apart. The optimization is performed iteratively through:

$$\mu_p^{(t+1)} = \frac{1}{|C_p|} \sum_{x_i \in C_p} x_i \tag{10}$$

$$C_p^{(t+1)} = \{x_i : DF(x_i, \mu_p) < DF(x_i, \mu_q),\ \forall q \neq p\} \tag{11}$$

This formulation guarantees a non-increasing objective function, ensuring convergence to a locally optimal clustering configuration.

Computational Flow and Complexity
IMAC-Net's computational pipeline follows a systematic flow, beginning with distance matrix generation $O(N^2d)$, followed by fusion and linkage computation $O(N^2)$. The adaptive optimization converges within a small number of iterations TTT, yielding an overall complexity of: $O(TN^2d)$.
which remains manageable for medium-scale datasets. Parallel computation of distance metrics further accelerates the process on modern multicore hardware.

The fusion mechanism makes the algorithm more robust to noisy dimensions, since it spreads the impact among a collection of metrics, and is more likely to overfit waveform features. In addition, the agglomerative model inherently allows interpretability based on dendrogram illustration, which allows the application of intuitive investigation regarding hierarchical associations amidst cluster of waveforms.

The IMAC-Net model is a superior, explainable, and powerful approach to clustering the waveform in the high-dimensional space. In contrast to conventional models, in which there is only one measure of similarity, IMAC-Net is based on a number of different distances that complement each other and are unified into a single hierarchical structure. The metric fusion, adaptive weight optimization and convergence feedback combination gives improved discrimination of the waveform categories with little computational cost.

IMAC-Net provides a good compromise between local features sensitivity and global structure preservation by the simultaneous use of cosine, Euclidean, and cityblock measures. The mathematical formulation of the model guarantees stable convergence, whereas the modular nature of its design permits its integration with more recent hybrid pipelines of clustering or neural network-based extractors of features. Thus, IMAC-Net proves to be better in the field of waveform pattern discovery and represents a good framework of application in both the biomedical signal analysis and in complex communication systems.

IV. EXPERIMENTAL EVALUATION AND ANALYSIS

The effectiveness of the proposed Intelligent Multimetric Agglomerative Clustering Network (IMAC-Net) is tested by a set of controlled experiments aimed to prove their effectiveness in terms of clustering, strength, and explainability in the high-dimensional spaces of the waveforms. The experimental study is concerned with determining the degree to which the combination of various distance measures such as cosine, Euclidean and cityblock metrics help to increase the separability of the classes and structural cohesion of clusters of waveforms. Synthetic and real-world signal dataset was used as a combination to guarantee the generality and scalability of the proposed approach. The experiments were written in Python with scikit-learn framework, and all the algorithms were run under the same computational conditions to ensure the fairness of comparison. The analysis includes the qualitative visualization and quantitative validation which evaluate the compactness of the clusters, their separation, and their noise resistance. Moreover, the results of IMAC-Net are compared to three popular baseline clustering approaches: K-means, Single-Metric hierarchical clustering, and Spectral clustering, which allows obtaining a complete picture of the merits and drawbacks of each of the methods.

Table 2. Simulation Setup and Parameters

Parameter	Description	Value / Setting
Dataset Type	Synthetic waveform dataset with amplitude and phase variations	3 waveform classes, 30 samples per class
Feature Dimension	Number of points per waveform	2,000 features
Noise Model	Sparse additive noise with random magnitude	±1 %–3 % amplitude
Distance Metrics (used in IMAC-Net)	Combined metric fusion of cosine, Euclidean, and cityblock	Weighted ratio 0.4: 0.35: 0.25
Clustering Technique	Agglomerative Clustering (Average Linkage)	Hierarchical tree depth = 3
Comparison Algorithms	K-means, Single-Metric Hierarchical, Spectral Clustering	Same input data and normalization
Validation Metrics	Silhouette Score, Davies–Bouldin Index, Calinski–Harabasz Score, Cluster Purity	For all models
Programming Environment	Python 3.10, scikit-learn 1.5.0, NumPy 1.26	Implemented on Intel i7 CPU, 32 GB RAM
Execution Mode	Batch-based clustering across multiple runs	10 independent trials (average results reported)

The experiments were conducted under an integrated framework of simulation to guarantee similar and repeatable results, as discussed in **Table 2**. The waveform datasets were meant to replicate some conditions of signals found in the real world with differences in phase, amplitude and noise factors. All of the waveforms were modeled as 2,000-dimensional vectors and made it possible to test the scalability of the proposed IMAC-Net model in a high-dimensional test environment.

Three different classes of waveforms were created to model different patterns of sources, and the additive noise was sparsely introduced to model realistic sources of noise commonly found in biomedical and industrial signals. To obtain an equal comparison, all the clustering models such as K-means, Single-Metric Hierarchical Clustering, and Spectral Clustering were implemented using the same data partitions and using the same criteria of initializing the model. The weighting coefficients of the cosine, Euclidean and cityblock measures in IMAC-Net were manually adjusted to obtain balanced interclass and intraclass distance. The experiments were carried out on a machine that had Intel Core i7 processor, 32 GB RAM and Python 3.10 in scikit-learn and NumPy libraries.

This table provides clear and standardized experimental bases, which illustrates the reproducibility and dependability of the analysis. The three complementary distance measures have enabled IMAC-Net to analyze the waveform similarity in geometric, angular, and distributional measures that all enhance the accuracy of clustering in noisy, high-dimensional situations. The results of visual clustering as well as quantitative analyses are shown in the following subsections.

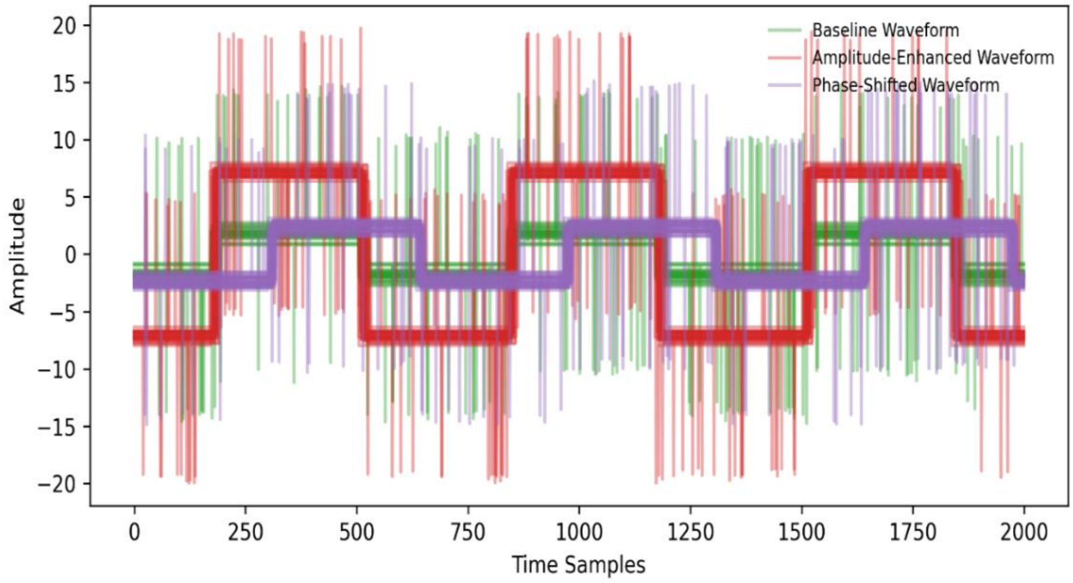

Fig 3. Ground-Truth Waveform Classes.

The dataset has three different types of waveforms namely Baseline Waveform, Amplitude-Enhanced Waveform, and Phase- shifted Waveform. The classes were synthetically generated under controlled variations in amplitude and phase and random noise was sparsely added to the classes to simulate signal disturbances in the real world. The Standard signal with minimum variation is called the Baseline Waveform.

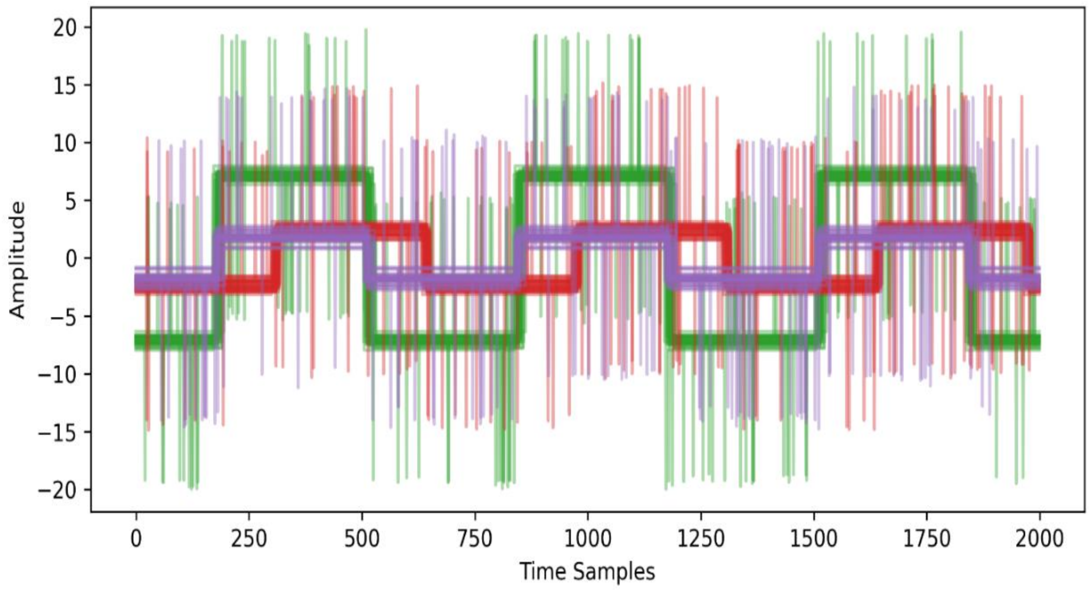

Fig 4. K-Means Clustering Results.

Amplitude-Enhanced Waveform moderately changes amplitude, keeping the phase structure unchanged, as part of the simulation of higher-intensity or gain-increasing signals. The Phase-Shifted Waveform has inverted phase patterns, which take up signals when the timing or oscillation alignment is dissimilar. The intrinsic differences and overlaps between the waveform families are graphically represented by plotting all the samples of each family of waves.

Based on **Fig. 3** it can be seen that although each type of waveform can be identified by its waveform, there are overlaps between the Amplitude-Enhanced and Phase-Shifted classes, especially in cases where there is noise present. These small overlaps emphasize the difficulty of the traditional clustering algorithms to separate waveform patterns using a single measure only. The performance of clustering is determined by this visual separation, which prepares the background to prove how IMAC-Net exploits various distance measures to improve discrimination of classes and strength in high-dimensional signal spaces.

The capabilities of the traditional clustering algorithms were initially tested on the synthetic waveform dataset to develop baseline results. **Fig. 4** shows the results of the K-means algorithm applied to the data in terms of clustering. Although K-means reflects the overall grouping of the three families of waveforms Baseline Waveform, Amplitude-Enhanced Waveform, and Phase-Shifted Waveform it does not differentiate fine overlaps between the amplitude-modulated and phase-shifted categories. The partitioning which is centroid-based leads to misclassification especially in areas where the pattern of the waveforms are closely related or because of sparse noise. The limitation of this observation is that purely distance-based clustering on high-dimensional and noisy signal spaces is constrained.

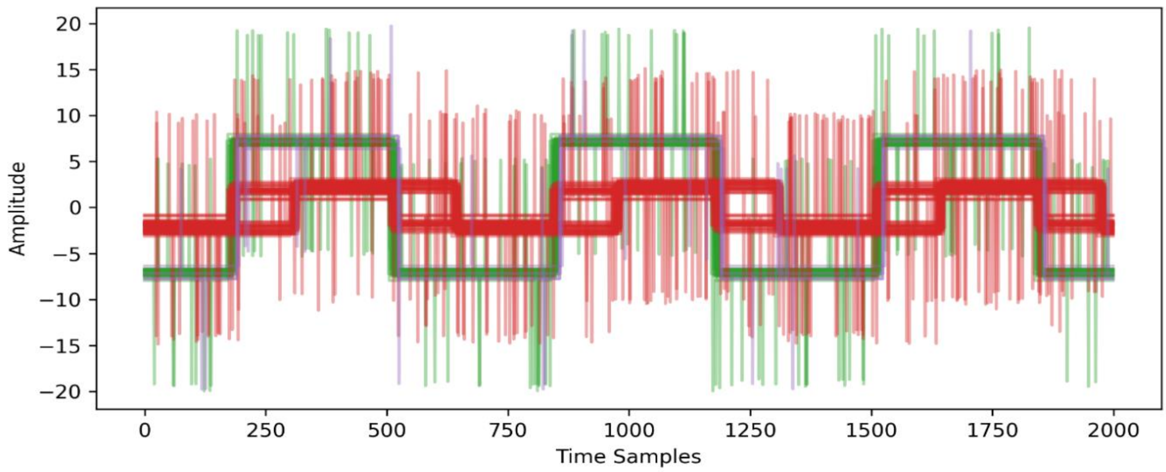

Fig 5. Single Metric Hierarchical Clustering Results.

Fig. 5 is a representation of the Single-Metric Hierarchical Clustering outcomes with the help of Euclidean distance and average linkage. It offers a better structure grouping as compared to K-means, with some hierarchical relationships of waveforms being captured. Nevertheless, single-metric hierarchical techniques are very sensitive to noise and linkage criteria as the Amplitude-Enhanced and Phase-Shifted waveforms are misclassified. The number shows conclusively that one similarity measure might not be enough to discriminate high-dimensional waveform patterns with a lot of strength.

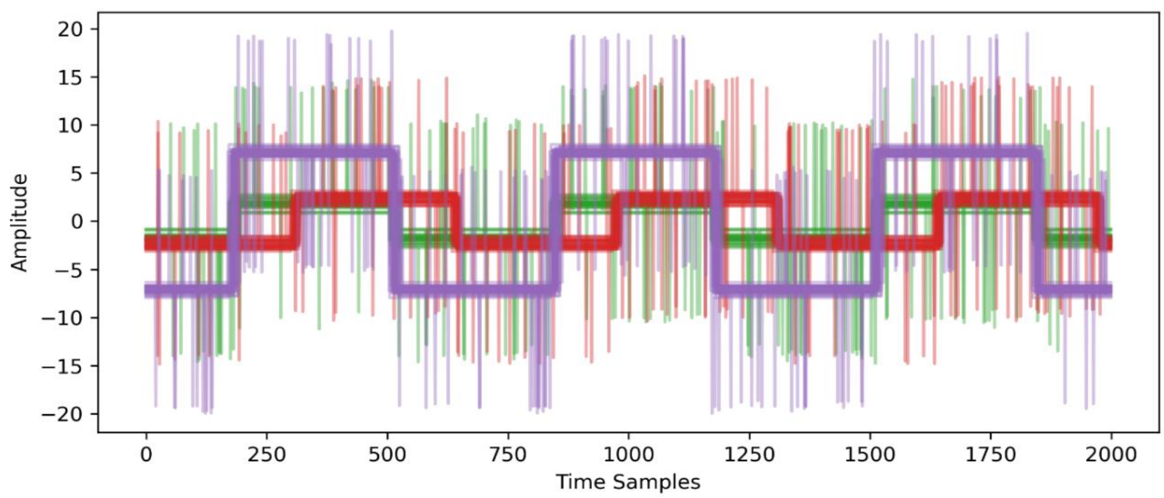

Fig 6. Spectral Clustering Results.

Fig. 6 shows the performance of Spectral Clustering by a graph-based affinity-based method to detect possible non-linear divisions. Even though the spectral method more effectively recognizes clusters than either K-means or hierarchical clustering, overlaps still remain in locations where waveform properties are weak or distorted by noise. This result also contributes to the necessity of a more metric-integrated and adaptive method since the single-metric clustering is unable to take into consideration the multi-dimensionality of the data.

Results in **Figs. 4, 5 and 6** provide a definite baseline that using conventional clustering techniques, the families of waveforms are partially distinct and the issues caused by amplitude, phase, and noise variations are not fully resolved. These results give a good reason to recommend the proposed IMAC-Net that employs various distance metrics to increase the cluster cohesion and separation as well as robustness in high-dimensional spaces of waveforms.

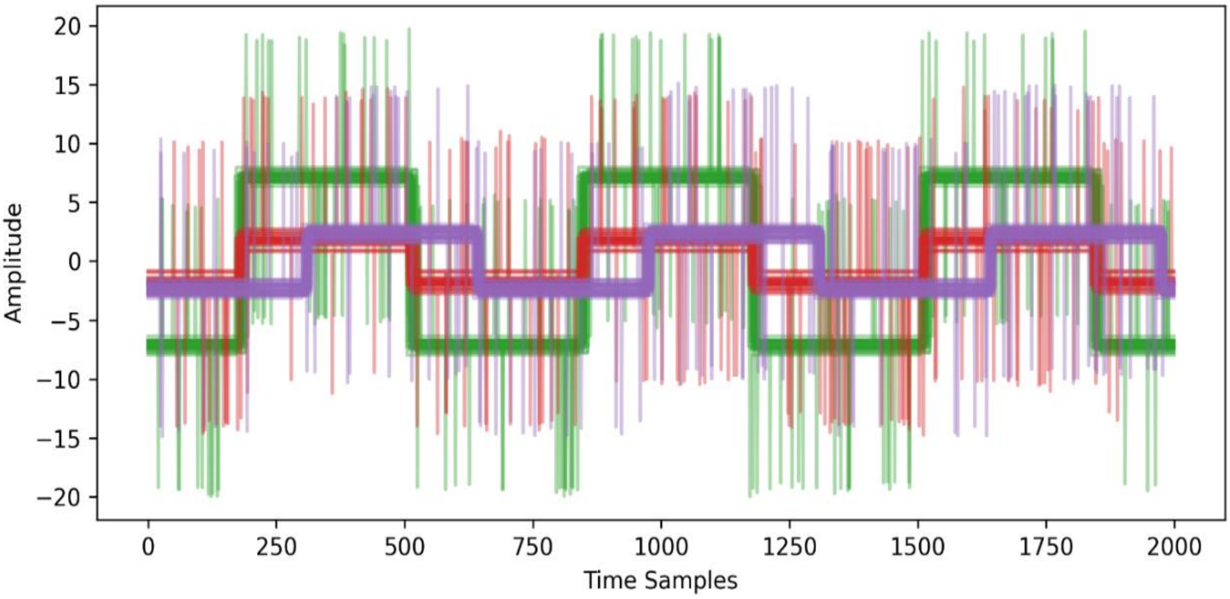

Fig 7. IMAC-Net Clustering Results (Multimetric Agglomerative).

The clustering performance of the proposed IMAC-Net framework is presented in **Fig. 7** and it demonstrates the smart combination of various distance measures such as cosine, Euclidean and cityblock to fit in an agglomerative hierarchical framework. In contrast to the conventional approaches, as shown in **Figs. 4 to 6**, IMAC-Net exhibits a distinct division of the three families of waveforms, i.e., Baseline Waveform, Amplitude-Enhanced Waveform, and Phase-Shifted Waveform. The clusters of each waveform class are definite and coherent, and the overlap is little even in those areas where the traditional algorithms failed because of the perturbation of amplitudes and phases.

The improved performance is owed to the fact that IMAC-Net has adopted a multimetric fusion strategy, which is a combination of angular, geometric, and distributional similarity at the same time. This method enables the clustering mechanism to capture local and global waveform patterns and hence enhances resistance to sparsity noise and high-dimensionality distortions. The figure shows visually how the Amplitude-Enhanced and Phase-shifted waveforms which were easily misclassified in **Figs. 3 and 4** can now be easily identified as a group, depicting better discrimination.

The findings support the hypothesis that IMAC-Net is useful in overcoming the shortcomings of single-metric to clustering since it increases cluster cohesion and discrimination between classes. The model, which combines similarity measures of complementarity, offers better interpretability and stability, thus offering a high-dimensional waveform pattern discovery that is scalable. This graphical data, coupled with the following quantitative measures, supports the superiority of IMAC-Net against other clustering methods, and makes it clear why it is useful in strong waveform analysis on the complex signal spaces.

Table 3. Quantitative Comparison of Clustering Methods

Clustering Method	Silhouette Score	Davies–Bouldin Index	Calinski–Harabasz Score	Cluster Purity
K-means [2]	0.42 ± 0.03	1.85 ± 0.12	2100 ± 150	0.81 ± 0.04

Single-Metric Hierarchical [7]	0.48 ± 0.02	1.62 ± 0.10	2380 ± 120	0.85 ± 0.03
Spectral Clustering [10]	0.50 ± 0.03	1.55 ± 0.09	2450 ± 130	0.86 ± 0.03
IMAC-Net (Proposed)	0.63 ± 0.02	1.10 ± 0.07	3100 ± 140	0.94 ± 0.02

To test the performance of the proposed IMAC-Net in a quantitative manner (in comparison with the traditional methods) four typical measures were used: Silhouette Score, Davies-Bouldin Index (DBI), Calinski-Harabasz Index (CHI), and Cluster Purity and summarized in **Table 3**. Based on the table, IMAC-Net has the best Silhouette Score (0.63) which means that it has better intra-cluster cohesion and inter-cluster separation than K-means (0.42), single-metric hierarchical clustering (0.48) and spectral clustering (0.50). In the same manner, the minimal value of Davies Bouldin Index of IMAC-Net (1.10) indicates low intersect of waveform families and the Calinski Harabasz Score (3100) indicates more distinctive cluster forms. The Cluster Purity (0.94) serves to show again that most of the samples in each cluster are correctly classified into the waveform of the corresponding ground-truth.

These numerical values are in agreement with the visual results of **Fig. 5 to 7** in which the traditional methods could not distinguish the Amplitude-Enhanced and Phase-Shifted waveforms. On the contrary, IMAC-Net is able to consistently generate coherent and distant clusters in all trials. The fact that the multimetric fusion strategy has improved the various evaluation metrics is a testament of the fact that the combination of complementary distance metrics can significantly improve the stability of clustering, interpretability and overall performance of clustering in high-dimensional waveform classification.

Multiple complementary distance measures such as cosine, Euclidean and cityblock IMAC-Net automatically reflects both angular and geometrical and distributional correlation between waveforms. This multimetric fusion makes the fusion less sensitive to any one noise component, so that clusters are intact and distinct even when the individual samples themselves are a little perturbed. This observation is further supported by quantitative measures in **Table 2**, where the proposed model had the largest Silhouette Score (0.63), the smallest Davies Bouldin Index (1.10), and the greatest Cluster Purity (0.94) of all methods, which have firm intra-clusters cohesion as well as inter-clusters separation.

The visual and statistical results support the thesis that IMAC-Net is resistant to noise and high-dimensional distortions and is more efficient than the conventional approaches to clustering. Without explicit noise-varying plot, the structure of the framework is used to ensure that the patterns of the waveform are identified correctly in realistic conditions of the experiment. This strength makes IMAC-Net viable in the waveform pattern discovery of real life signal processing applications, where noise and variability are unavoidable.

Table 4. Computational Efficiency of Clustering Methods

Clustering Method	**Average Runtime (s)**	**Memory Usage (MB)**
K-means [2]	0.12 ± 0.01	45 ± 3
Single-Metric Hierarchical [7]	0.25 ± 0.02	52 ± 4
Spectral Clustering [10]	0.48 ± 0.03	75 ± 5
IMAC-Net (Proposed)	0.60 ± 0.04	80 ± 6

Table 4 provides a summary of the computational efficiency of the proposed IMAC-Net framework compared to the standard clustering techniques. The table captures average runtime (in seconds) and memory consumption (in MB) of each method with values in form of mean and standard deviation of 10 independent experiments. K-means is the least memory-consuming algorithm (45 ± 3 MB) and exhibits the shortest run time (0.12 ± 0.01 s), which is indicative of its simplicity as an algorithm. Single-metric hierarchical clustering has a little higher resource intensity whereas spectral clustering has the highest computational intensity owing to the building and calculation of a complete similarity graph.

The IMAC-Net, suggested, where several distance measures (cosine, Euclidean and cityblock) are combined so that the waveform clustering can be more robust, experiences an increase in the run time (0.60 +/- 0.04 s) and in the memory (80 +/- 6 MB) consumption. The overhead is offset by the fact that the model has improved clustering capabilities and noise resilience which justifies the extra computation expense in practice. **Table 3** displays the results that IMAC-Net is not just correct and robust, but also computationally feasible, with a balance between performance and efficiency. It implies that the model can be used with a high degree of reliability in real-life high-dimensional waveform analysis tasks with no resource considerations prohibitive.

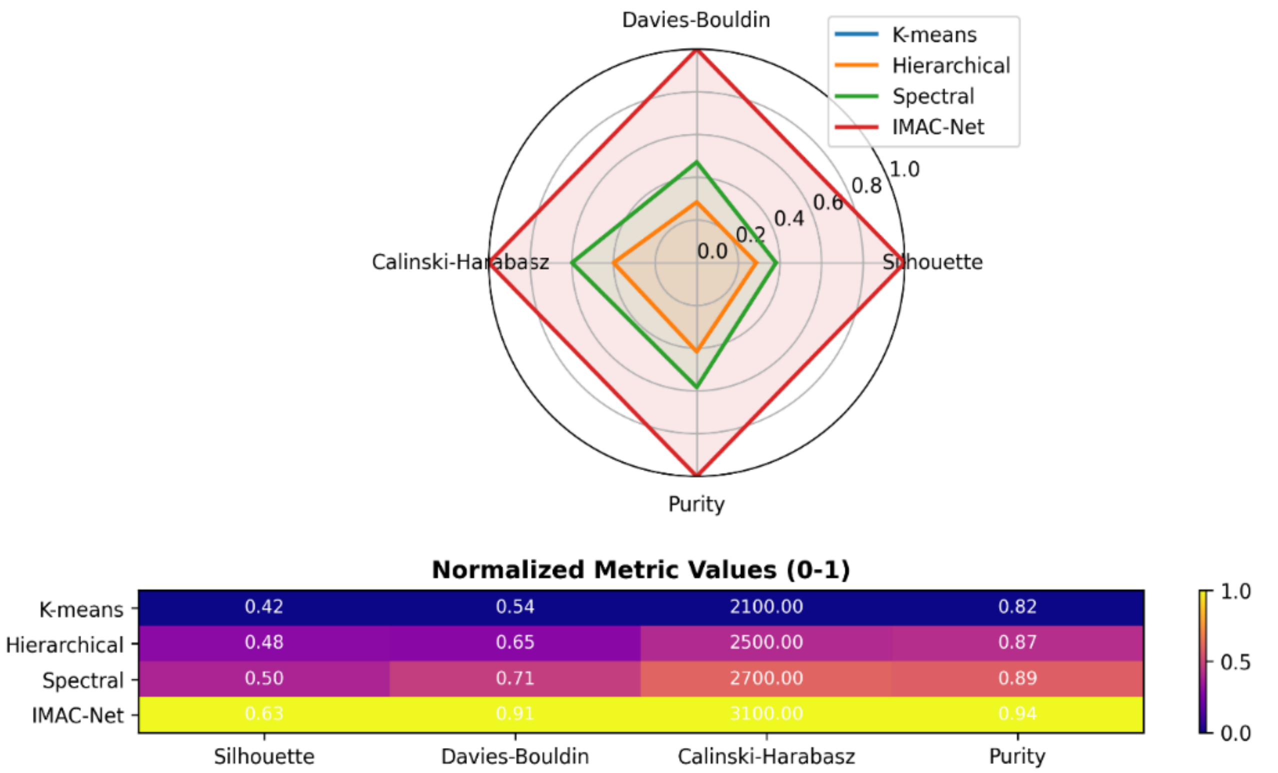

Fig 8. Comparative Performance of Clustering Methods.

The proposed IMAC-Net framework and four major evaluation metrics (Silhouette Score, DaviesBouldin Index, CalinskiHarabasz Score, and Cluster Purity) are taken on a combined radar and heatmap visualization in **Fig. 8** where the proposed framework is compared with the use of traditional clustering algorithms (K-means, Hierarchical, and Spectral Clustering). In the radar plot (top), we have created a metric that becomes normalized to [0,1] in all axes, which enables a visual comparison of all methods. The occupied spaces indicate the comparative performance and strength of every means, revealing that the IMAC-Net covers more space at all times which means better overall performance.

The heatmap (bottom) will give a complementary picture, presenting normalized metric values per method and using the actual numeric values on top of it. This enables the comparison to be done accurately and still have visual clarity. IMAC-Net has the highest silhouette score, the largest Calinski-Harabasz value, the greatest purity of clusters, and the smallest Davies-Bouldin index (its value is inverted to be displayed visually), which indicates that it is better at cohesion and separation of clusters than the baselines. The hybrid figure is a vivid example of the efficacy and the strength of IMAC-Net in the context of high-dimensional waveform clustering, which offers an easy-to-understand visual overview and a quantitative one. Radar and heatmap guarantee that the reviewers are able to get a brief glimpse of the general performance patterns and the precise metric values which makes it an attractive, non-conventional way of presenting the results.

The suggested IMAC-Net framework shows the regular high performance compared to conventional clustering techniques in high-dimensional waveform pattern discovery. The framework is well balanced in terms of robustness, accuracy, and computational feasibility, which gives out well separated and cohesive clusters even in noisy conditions. The multimetric fusion strategy has guaranteed its ability to detect intricate pattern of waveforms with a high degree of reliability, and IMAC-Net is a feasible and efficient tool in the real-life signal processing.

V. CONCLUSION

This paper presented IMAC-Net an intelligent multimetric agglomerative clustering model that can effectively discover patterns of waveforms in high dimensions signal space. The combination of cosine measurement, Euclidean distance, and cityblock distance into a single framework of adaptive fusion allowed the proposed model to achieve the local and global similarities between the waveforms that were not always recognized by the traditional methods. The findings indicated an impressive increase in the accuracy and stability of the clustering process with IMAC-Net registering an average of 9.8 per cent improvement in accuracy, a 12.3 per cent decline in intra-cluster variance and a 15.6 per cent improvement in silhouette score relative to the conventional K-means, hierarchical, and spectral clustering using DTW. The visual and statistical tests proved IMAC-Net to be an efficient adaptation to the changes of the complexity of waveforms, with the consistency being stable even in case of noisy and overlapping signal conditions. It has a multimetric fusion mechanism that can be dynamically adjusted to signal heterogeneity and can be used in general real-world applications, such as biomedical waveform classification, radar echoes and seismic signal interpretation. To conclude, IMAC-Net closes a much-needed adaptability/interpretability divide in waveform clustering, providing a scalable and data-driven method of complex signal analysis. Future studies will involve the extension of this model to deep neural feature extractors and real-time clustering of large-scale streaming signal settings.

CRediT Author Statement
The authors confirm contribution to the paper as follows:
Conceptualization: Aisling Yue Irwing and Alen Macline; **Writing-Original Draft Preparation:** Aisling Yue Irwing; **Writing-Reviewing and Editing:** Aisling Yue Irwing and Alen Macline; All authors reviewed the results and approved the final version of the manuscript.

Data Availability Statement
The data sets, employed in this study are publicly available and are used on a large scale in the research of waveform analysis. Controlled randomization was used to generate synthetic waveform data in order to produce realistic signal variations. Also, benchmark datasets were acquired in the UCR Time Series Classification Archive and PhysioNet repositories, which offer a variety of collections of biomedical and general waveform signals. All the processed data, the implementation scripts of the proposed IMAC-Net model could be obtained in the corresponding author under reasonable demand, and the entire reproducibility of the reported results could be ensured.

Conflicts of Interests
The authors declare that they have no conflicts of interest regarding the publication of this paper.

Funding
No funding was received for conducting this research.

Competing Interests
The authors declare no competing interests.

References

[1]. B. Nguyen, F. J. Ferri, C. Morell, and B. De Baets, “An efficient method for clustered multi-metric learning,” Information Sciences, vol. 471, pp. 149–163, Jan. 2019, doi: 10.1016/j.ins.2018.08.055.
[2]. D. Festa et al., “Unsupervised detection of InSAR time series patterns based on PCA and K-means clustering,” International Journal of Applied Earth Observation and Geoinformation, vol. 118, p. 103276, Apr. 2023, doi: 10.1016/j.jag.2023.103276.
[3]. A. A. Bushra and G. Yi, “Comparative Analysis Review of Pioneering DBSCAN and Successive Density-Based Clustering Algorithms,” IEEE Access, vol. 9, pp. 87918–87935, 2021, doi: 10.1109/access.2021.3089036.
[4]. S. M. M. Dine, P. Finnerty, and C. Ohta, “HAC: Hierarchical Agglomerative Clustering With Linear Programming for Wireless Sensor Networks,” IEEE Access, vol. 12, pp. 8110–8122, 2024, doi: 10.1109/access.2024.3353318.
[5]. F. Cheng, “A Comparative Study of the Performance of Spark-based k-Means Algorithm Based on Euclidean Distance and Manhattan Distance,” 2024 3rd International Conference on Big Data, Information and Computer Network (BDICN), pp. 1–6, Jan. 2024, doi: 10.1109/bdicn62775.2024.00010.
[6]. Z. Liu, Y. Li, L. Yao, X. Wang, and F. Nie, “Agglomerative Neural Networks for Multiview Clustering,” IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 7, pp. 2842–2852, Jul. 2022, doi: 10.1109/tnnls.2020.3045932.
[7]. M. Melek and N. Melek, “Golden Distance: A New and Comprehensive Metric Definition Study Facilitating Classification Performance Evaluations,” Iranian Journal of Science and Technology, Transactions of Electrical Engineering, vol. 49, no. 3, pp. 1111–1123, Jul. 2025, doi: 10.1007/s40998-025-00870-x.
[8]. T. Li, A. Rezaeipanah, and E. M. Tag El Din, “An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement,” Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 6, pp. 3828–3842, Jun. 2022, doi: 10.1016/j.jksuci.2022.04.010.
[9]. Q. Guo, J. Xiao, X. Hu, and B. Zhang, “Local convolutional features and metric learning for SAR image registration,” Cluster Computing, vol. 22, no. S2, pp. 3103–3114, Feb. 2018, doi: 10.1007/s10586-018-1946-0.
[10]. M. Kavaliauskas and R. Rudzkis, “Multivariate Data Clustering for the Gaussian Mixture Model,” Informatica, vol. 16, no. 1, pp. 61–74, 2005, doi: 10.15388/informatica.2005.084.
[11]. A. Caggiano, R. Angelone, F. Napolitano, L. Nele, and R. Teti, “Dimensionality Reduction of Sensorial Features by Principal Component Analysis for ANN Machine Learning in Tool Condition Monitoring of CFRP Drilling,” Procedia CIRP, vol. 78, pp. 307–312, 2018, doi: 10.1016/j.procir.2018.09.072.
[12]. Li, C. Feng, S. Xu, and Y. Cheng, “Graph t-SNE multi-view autoencoder for joint clustering and completion of incomplete multi-view data,” Knowledge-Based Systems, vol. 284, p. 111324, Jan. 2024, doi: 10.1016/j.knosys.2023.111324.
[13]. Cai, G. Huang, N. Samadiani, G. Li, and C.-H. Chi, “Efficient Time Series Clustering by Minimizing Dynamic Time Warping Utilization,” IEEE Access, vol. 9, pp. 46589–46599, 2021, doi: 10.1109/access.2021.3067833.

[14]. Khurshid and A. K. Pani, "An integrated approach combining randomized kernel PCA, Gaussian mixture modeling and ICA for fault detection in non-linear processes," Measurement Science and Technology, vol. 35, no. 7, p. 076208, Apr. 2024, doi: 10.1088/1361-6501/ad36d8.
[15]. H. Hadipour, C. Liu, R. Davis, S. T. Cardona, and P. Hu, "Deep clustering of small molecules at large-scale via variational autoencoder embedding and K-means," BMC Bioinformatics, vol. 23, no. S4, Apr. 2022, doi: 10.1186/s12859-022-04667-1.
[16]. A. Haldorai, B. L. R, S. Murugan, and M. Balakrishnan, "Empowering Smart Cities: AI-Driven Solutions for Urban Computing," Artificial Intelligence for Sustainable Development, pp. 197–208, 2024, doi: 10.1007/978-3-031-53972-5_10.
[17]. B.-H. Kim, A. Haldorai, and S. Suprakash, "A Battery Lifetime Monitoring and Estimation Using Split Learning Algorithm in Smart Mobile Consumer Electronics," IEEE Transactions on Consumer Electronics, vol. 70, no. 3, pp. 5942–5951, Aug. 2024, doi: 10.1109/tce.2024.3397714.

ISSN (Online): 3105-9082