A Hybrid Collaborative Clustering Approach for Noise-Robust Speech Recognition

Ameni Filali
Department of Computer Engineering,
College of Computer Science and Engineering,
University of Hail 81481, Saudi Arabia

ABSTRACT

Collaborative clustering is an emerging field in machine learning that reveals the common structure within relevant and noisy datasets distributed across different sites. The value of collaborative clustering lies in its wide range of potential applications, including multi-view clustering and knowledge transfer. In this paper, a hybrid collaborative clustering approach is proposed that combines both horizontal and vertical collaborative clustering using AdSOM, which is efficient for all prototype-based clustering methods. The benefit of collaboration between datasets is quantified using a collaboration coefficient, which is evaluated iteratively during the collaboration step. This process is optimized using the steepest descent method.

To demonstrate the effectiveness of the proposed collaborative approaches, a case study is conducted on phoneme recognition in continuous speech and in a speaker-independent context. The collaborative approaches are validated using two datasets: TIMIT and NTIMIT. Experimental results show promising performance.

General Terms

Speech Recognition, Machine Learning, Self-organization Map.

Keywords

Unsupervised learning; Collaborative clustering; SOM based on a locally adapting neighborhood radii (AdSOM); Random Subspaces Method (RSM).

1. INTRODUCTION

Human listeners are able to recognize speech in noisy environments, whereas most classic speech recognition methods do not perform well in the presence of noise. Unlike conventional Mel-frequency cepstral coefficient (MFCC)-based method, this work suggests a phoneme classification technique using the collaborative approach. The features of neurograms were used to learn the recognition system using a SOM based on a locally adapting neighborhood radii clustering method. Performance was evaluated for different types of phonemes such as stops, vowels from the TIMIT dataset for both under quiet and noisy conditions (NTIMIT). This method could be used in the field of speech recognition such as speech to text application, especially under noisy conditions. In this paper, it is assumed that a group of datasets is distributed across different sites; the data may describe remote sensing images, phonemes, medical images, hyperspectral imaging, etc. In the literature, several studies have been proposed that enable collaboration between different distributed datasets across multiple centers while preserving data confidentiality ([1], [2], [3], [4], [5]). The collaborative approach was introduced in the context of dimensionality reduction. However, in the case of very large datasets where all features are relevant and no reduction can be performed, the data can be partitioned into subsets. Collaborative clustering can then be applied to cluster the dataset. This paper is an extension of work originally presented in International Conference on Computer Systems and Applications (AICCSA) [6]. Collaborative clustering was first introduced by Pedrycz[7], using a fuzzy k-means algorithm. The fundamental concept of collaboration is: “the clustering algorithms operate locally but collaborate by exchanging information about their findings” Pedrycz. There are three principal types of collaborative datasets: vertical, horizontal, and hybrid collaboration. In this study, the focus is on vertical and horizontal topological collaborative clustering, and a new approach is proposed that combines these two topological collaborative paradigms (vertical and horizontal) using SOM based on a locally adapting neighborhood radius (AdSOM). In the case of the vertical collaboration approach, the data sets could contain data about different observations represented by the same features. Otherwise, In the case of the horizontal collaboration approach, the datasets could represent the same observations but with different variables. This case is the most complicated one, since different feature spaces mean that samples are illustrated in different variables, thus different dimensions.

This approach can be treated as multi-view clustering where the processing is done on multi-represented data. Therefore, in this case, the horizontal collaboration approach is presented. The collaborative process is divided into two steps: a local step and a collaborative step. The local step performs clustering on each dataset based on reference vectors, producing a score for each dataset. The collaborative step then enables each dataset to collaborate with the clustering results obtained from the other datasets during the local step. In this study, the collaboration step is divided into two phases: the first involves computing the reference matrix, while the second determines confidence connections for clustering.

In this article, a collaborative clustering approach based on AdSOM is described as the local clustering step. The collaboration coefficients are estimated using the steepest descent method [21] inspired by ([8], [9]).

can also be estimated using alternative methods, such as Particle Swarm Optimization (PSO) [10] driven algorithm in ([8],[11],[12]), the multi-PSO proposed in [13] and the gradient method used in the work of Ghassany[9].

In addition, another method for estimating the coefficient of collaboration automatically is introduced in [14], based on a similarity measure of partition matrices. But this approach operates just for horizontal collaboration.

The aim of speech recognition systems is to achieve the best possible clustering performance. Different classifiers usually make errors on different samples; therefore, combining classifiers can lead to more accurate recognition rates. Lately in the domain of machine learning, the concept of combining classifiers is suggested as a new direction for performance's improvement of individual classifiers. This approach is seen as the set of classifiers (EOC) [15]. Diverse classifiers can be established in several ways, such as Random Subspaces [16, 17], Bagging and Boosting [18]. Besides, the classifier variety does not occur randomly but is produced systematically by various ensemble creation methods. When using diverse data subsets to form classifiers, these methods may create various classifiers for the Ensemble of Classifiers. In this work, a program is presented to measure data diversity directly from random subspaces and to identify the most suitable subsets of data for training an ensemble of classifiers.

Random Subspaces is an ensemble classifier that creates diverse classifiers each operating in a subspace of the original feature space by using various feature subsets to train them. All these ensemble creation approaches assume benefit of several data, which are assembled by extracting only a subset of variables or a subset of samples, thus creating diverse classifiers. Random subspace approach has been applied for decision trees, [18] linear classifiers, support vector machines, nearest neighbors and other types of classifiers.

Many studies have shown that multiple classifier systems, such as the random subspace method (RSM), are more robust than a single classifier and give the most remarkable results for many important speech recognition issues.

In this paper, a comparative study is conducted between horizontal, vertical, and hybrid collaborative clustering approaches, as well as the ensemble method known as the Random Subspace Method.

In this paper, the performance of these approaches is assessed on the TIMIT and NTIMIT datasets using several validation criteria. The results, which are promising, are reported in tables and figures.

This paper is organized as follows: Section 2 provides a brief introduction to the Self-Organizing Maps algorithm based on a locally adapting neighborhood radius (AdSOM). Section 3 introduces the principle of collaborative clustering. The methodologies for the collaborative approaches (vertical, horizontal, and hybrid) are presented in Sections 3.1, 3.2, and 3.3. Section 3.4 provides an overview of randomization methods for constructing ensembles of classifiers. The experimental results and discussion are presented in Section 4, and Section 5 concludes the paper with perspectives for future work.

2. ADSOM AS LOCAL PHASE FOR COLLABORATIVE CLUSTERING

The Self-Organizing Maps (SOM) is suggested by T. Kohonen [19] have been broadly applied for clustering and visualization of multidimensional datasets. Therefore, there are several varieties of derivative methods for topological maps from the original model proposed by Kohonen ([19], [20], [21], [22]).
To minimize the topological error Kiviluoto [19] has proposed a variant SOM called AdSOM. In this variant each neuron i has its own neighborhood width parameter σ_i attached with it. During learning of the AdSOM, the parameter σ_i is collection to half the diameter of the lattice for all i. During learning, input vectors are specified to the AdSOM as performed already to the SOM and weights are adjusted according to the update rule while minimizing learning rate. In addition, the neighborhood width is defined by the BMU value σ_i. The parameter σ_i is presented by the local topographic errors. In fact, a local topographic error occurs when an input vector x, let w_j and w_k are the two closest weighting vectors and the corresponding BMU w_j and w_k are not adjacent. Moreover, for units w_i that are close to the two BMUs, a new value for the neighborhood radius σ_i is determined [1].

$$\sigma_i = \begin{cases} \|w_j - w_k\| & if\ max\{\|w_i - w_j\|, \|w_i - w_k\|\} \le \|w_j - w_k\| \\ \|w_j - w_k\| - s & when\ s = min\{\|w_i - w_j\|, \|w_i - w_k\|\} < \|w_j - w_k\| \\ 1 & Otherwise. \end{cases}$$

3. COLLABORATIVE CLUSTERING

In the context of collaborative classification and according to the structure of the dataset to work, there are three principal types of collaboration: vertical, horizontal and hybrid collaborations ([23], [24], [25]). In this work, particular interest is given to horizontal and vertical collaboration. A combination of both approaches is also proposed to obtain a hybrid collaborative clustering method. This paper studies the collaboration between several unsupervised classifications from self-organizing maps. Each dataset is classified using the AdSOM approach. To simplify the analysis, the maps of the different datasets have the same dimensions (number of neurons) and the same structure (topology).

3.1 Unsupervised horizontal collaborative clustering

In the case of horizontal collaborative clustering, all datasets define the same observations. So, all these collaborative datasets have the same number of observations, but in different spaces of description ([26], [27]).
The main idea of the founding principle of horizontal collaboration [8] between different maps AdSOM is that if an individual of the ii-th dataset is projected on the j-th neuron of the ii-th map AdSOM, then that same individual in jj-th dataset will be projected on the same neuron j of the jj-th map or one of its neighboring neurons. Moreover, the neurons corresponding to different maps must capture the same observations. For this reason, an additional term reflecting the principle of collaboration is added to the AdSOM objective function. The function is weighted using a collaboration parameter that represents the degree of collaboration. The new objective function is determined by two terms:

$$R_H^{[ii]}(\chi, w) = R_{SOM}^{[ii]}(\chi, w) + R_{ColH}^{[ii]}(\chi, w) \qquad (1)$$

$$R_{SOM}^{[ii]}(\chi, w) = \sum_{i=1}^{N} \sum_{j=1}^{|w|} K_{\sigma(j,\chi(x_i))}^{[ii]} \left\|x_i^{[ii]} - w_j^{[ii]}\right\|^2 \qquad (2)$$

$$R_{ColH}^{[ii]}(\chi, w) = \sum_{jj=1, jj\neq ii}^{P} \alpha_{[ii]}^{[jj]} \sum_{i=1}^{N} \sum_{j=1}^{|w|} \left(K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]}\right)^2 \left\|x_i^{[ii]} - w_j^{[ii]}\right\|^2 \qquad (3)$$

Where P represents the number of datasets, N represents the number of individuals; |w| represents the number of prototype vectors from the ii-th AdSOM map. $\chi(x_i)$ represents the function of assignment for discovering the winner neuron (BMU) it selects the neuron with the nearest referent from the data x_i involving the Euclidean distance of the given x_i:

$$\chi(x_i) = argmin(\|x_i - w_j\|^2) \qquad (4)$$

σ(i, j) represents the distance between two neurons i and j of the AdSOM map. Kσ(i,j) represents the neighborhood function on the AdSOM map[cc]. Kσ(i,j) depends on the distance between two neurons and is calculated as follows:

$$K_{\sigma(i,j)}^{[cc]} = exp\left(-\frac{\sigma^2(i,j)}{\lambda^2(t)}\right) \qquad (5)$$

Where λ (t) represents the temperature function modeling the extent of the neighborhood:

$$\lambda(t) = \lambda_i(\frac{\lambda_f}{\lambda i})^{\frac{t}{T_{max}}} \qquad (6)$$

Where λ_i and λ_f are respectively the initial temperature and the final temperature (e.g. $\lambda_i = 2$ and $= \lambda_f$ 0, 5) and T_{max} is the maximum number allocated to the time (number of iterations x the number of training examples).

The algorithm of horizontal collaborative clustering is presented in algorithm1. This method can be splitted into two steps: a local step and a collaboration step. The local step allows applying the enriched algorithm AdSOM, separately and locally on each of the databases. The collaboration step would cooperate every of the databases with all AdSOM maps associated with other bases obtained after the local step.

3.2 Unsupervised vertical collaborative clustering

In the case of vertical collaboration, all databases have the same features but have different individuals [9]. In this case, individuals of these databases have the same size, and the size referents vectors for all AdSOM maps will be the same. The basic idea of vertical collaboration is to ensure that neuron j of the ii-th SOM map and the same neuron j of the jj-th map are close in terms of Euclidean distance. Otherwise, neurons belonging to different maps should represent similar groups of observations.

To enforce this constraint during the collaborative phase, an additional term is added to the classical AdSOM objective function across the different maps. Formally, the following new objective function is obtained:

$$R_V^{[ii]}(\chi, w) = R_{SOM}^{[ii]}(\chi, w) + R_{ColV}^{[ii]}(\chi, w) \qquad (7)$$

$$R_{ColV}^{[ii]}(\chi, w) = \sum_{jj=1, jj\neq ii}^{P} \alpha_{[ii]}^{[jj]} \sum_{i=1}^{N[ii]} \sum_{j=1}^{|w|} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2 \left\| w_j^{[ii]} - w_j^{[jj]} \right\|^2 \qquad (8)$$

Where P represents the number of datasets, N[ii] represents the number of individuals of the ii-th database; |w| represents the number of vectors of the referents AdSOM map ii which is the same for all maps. The equation (8) is minimized for each dataset in collaboration step.

The vertical collaborative clustering method using self-organizing maps [32] is illustrated in Algorithm2 including two steps: the local step and collaboration step. As for horizontal collaboration, vertical will begin immediately with the collaboration step if the maps are previously known.

3.3 Hybrid Approach

In this section, the proposed contribution is to combine both topological collaborative approaches: Horizontal and Vertical collaboration presented in the previous section using Self-Organizing Map based on a locally adapting neighborhood radii. In the case of hybrid collaborative clustering, all datasets describe the different observations and the different features. Then, all these collaborative databases have different numbers of observations, also in different spaces of description: the different number of individuals and a different number of variables.

The principal idea of the founding principle of hybrid collaboration between different AdSOM maps is that if an individual of the ii-th dataset is projected on the j-th neuron of the ii-th map AdSOM, afterward, that same individual in jj-th dataset will be projected on the same neuron j of the jj-th map or one of its neighboring neurons. The main purpose of hybrid collaboration is to establish a grid model of horizontal and vertical clustering.

Besides horizontal and vertical clustering, which represent two general models of collaboration, a variety of intermediate situations can be considered, where patterns from different sources are grouped into common subsets of the dataset and assigned to the same variable. Also, one can achieve mechanisms of horizontal and vertical collaboration at the same time. Therefore, a grid-based clustering mode is obtained, and examples of collaboration are illustrated in Figure 3. The objective function proposed for the ii-th pattern as a subject of minimization is an aggregation of the components used in the previous models of collaborative clustering. Knowing that the summation points at the corresponding datasets, which is, $\mathbf{R_{colH}}$ involves all datasets that effectuate in the horizontal method of clustering, whereas $\mathbf{R_{colV}}$ concerns those using vertical collaboration. There are also many hybrid methods of collaboration implicating datasets with possible connections of vertical and horizontal collaboration. The new objective function is composed of two terms:

$$R_{Col_{Hybrid}}^{[ii]}(\chi, \omega) = R_{Col_H}^{[ii]}(\chi, \omega) + R_{Col_V}^{[ii]}(\chi, \omega) \qquad (9)$$

Algorithm1: Horizontal collaboration

```
Random the collaboration coefficient $\alpha^{[jj]}_{[ii]}$
1. Local step
For each BD[ii], ii = 1 à P
Find the referents minimizing the enriched AdSOM:
    $w^* = \underset{w}{argmin} \left[ R_{SOM}^{[ii]}(\chi, \omega) \right]$
2. Collaboration step
For the horizontal collaborative clustering of the[ii] map with [jj] map:
Update the referents of the [ii] map minimizing the objective function of the topological horizontal collaboration involving the expression:
    $w^{*[ii]} = \underset{w}{argmin} \left[ R_H^{[ii]}(\chi, \omega) \right] = \underset{w}{argmin} \left[ R_{SOM}^{[ii]}(\chi, \omega) + R_{Col_H}^{[ii]}(\chi, \omega) \right]$
With:
```

$$w_{jk}^{*[ii]}(t+1) = w_{jk}^{*[ii]}(t) + \frac{\sum_{i=1}^{N} K_{\sigma(j,\chi(x_i))}^{[ii]} x_{ik}^{[ii]} + \sum_{jj=1,jj\neq ii}^{P} \sum_{i=1}^{N[ii]} \alpha_{[ii]}^{[jj]} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2 \, x_{ik}^{[ii]}}{\sum_{i=1}^{N} K_{\sigma(j,\chi(x_i))}^{[ii]} + \sum_{jj=1,jj\neq ii}^{P} \sum_{i=1}^{N} \alpha_{[ii]}^{[jj]} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2}$$

Update the collaboration coefficients using the Steepest Descent Method:

$$\alpha_{[ii]}^{[jj]}(t+1) = \alpha_{[ii]}^{[jj]}(t) + \frac{\sum_{i=1}^{N} \sum_{j=1}^{|w|} K_{\sigma(j,\chi(x_i))}^{[ii]}}{2\sum_{i=1}^{N} \sum_{j=1}^{|w|} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2}$$

Algorithm2: vertical collaboration

Random the collaboration coefficient $\alpha^{[jj]}_{[ii]}$

1. Local step

For each BD[ii], ii = 1 à P

Find the referents minimizing the enriched AdSOM:

$$w^* = \underset{w}{argmin}\left[R_{SOM}^{[ii]}(\chi,\omega)\right]$$

2. Collaboration step

For the vertical collaborative clustering of the[ii] map with [jj] map:

Update the referents of the [ii] map minimizing the objective function of the vertical collaboration involving the expression:

$$w^{*[ii]} = \underset{w}{argmin}\left[R_V^{[ii]}(\chi,\omega)\right] = \underset{w}{argmin}\left[R_{SOM}^{[ii]}(\chi,\omega) + R_{Col_V}^{[ii]}(\chi,\omega)\right]$$

with:

$$w_{jk}^{*[ii]}(t+1) = w_{jk}^{*[ii]}(t) + \frac{\sum_{i=1}^{N} K_{\sigma(j,\chi(x_i))}^{[ii]} x_{ik}^{[ii]} + \sum_{jj=1,jj\neq ii}^{P} \sum_{i=1}^{N} \alpha_{[ii]}^{[jj]} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2 \, w_{ik}^{[jj]}}{\sum_{i=1}^{N} K_{\sigma(j,\chi(x_i))}^{[ii]} + \sum_{jj=1,jj\neq ii}^{P} \sum_{i=1}^{N} \alpha_{[ii]}^{[jj]} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2}$$

Update the collaboration coefficients applying the Steepest Descent Method:

$$\alpha_{[ii]}^{[jj]}(t+1) = \alpha_{[ii]}^{[jj]}(t) + \frac{\sum_{i=1}^{N} \sum_{j=1}^{|w|} K_{\sigma(j,\chi(x_i))}^{[ii]} \left\| x_i^{[ii]} - w_j^{[ii]} \right\|^2}{2\sum_{i=1}^{N} \sum_{j=1}^{|w|} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2 \left\| w_j^{[ii]} - w_j^{[jj]} \right\|^2}$$

Algorithm3: Topological hybrid collaboration

Random the collaboration coefficient $\alpha^{[jj]}_{[ii]}$

1. Local step

For each BD[ii], ii = 1 à P

Find the referents minimizing the enriched AdSOM :

$$w^* = \underset{w}{argmin}\left[R_{SOM}^{[ii]}(\chi,\omega)\right]$$

2. Collaboration step

For the hybrid collaboration of the[ii] map(horizontal) with [jj] map(vertical):

Update the referents of the [ii] map minimizing the objective function of the hybrid collaboration involving the expression:

$$w^{*[ii]} = \underset{w}{argmin}\left[R_{Col_{Hybrid}}^{[ii]}(\chi,\omega)\right] = \underset{w}{argmin}\left[R_{Col_H}^{[ii]}(\chi,\omega) + R_{Col_V}^{[ii]}(\chi,\omega)\right]$$

With:

$$w_{jk}^{*[ii]}(t+1) = w_{jk}^{*[ii]}(t) + \frac{\sum_{i=1}^{N} K_{\sigma(j,\chi(x_i))}^{[ii]} x_{ik}^{[ii]} + \sum_{jj=1,jj\neq ii}^{P} \sum_{i=1}^{N[ii]} \alpha_{[ii]}^{[jj]} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2 \, (x_{ik}^{[ii]} + w_{ik}^{[jj]})}{\sum_{i=1}^{N} K_{\sigma(j,\chi(x_i))}^{[ii]} + \sum_{jj=1,jj\neq ii}^{P} \sum_{i=1}^{N} \alpha_{[ii]}^{[jj]} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2}$$

Update the collaboration coefficients applying the Steepest Descent Method:

$$\alpha_{[ii]}^{[jj]}(t+1) = \alpha_{[ii]}^{[jj]}(t) + \frac{\sum_{i=1}^{N} \sum_{j=1}^{|w|} K_{\sigma(j,\chi(x_i))}^{[ii]} w_j^{[ii]}}{2\sum_{i=1}^{N} \sum_{j=1}^{|w|} (K_{\sigma(j,\chi(x_i))}^{[ii]} - K_{\sigma(j,\chi(x_i))}^{[jj]})^2}$$

$$R_{Col_{Hybrid}}^{[ii]}(\chi,\omega) = \sum_{jj=1,jj\neq 1}^{P} \alpha_{[ii]}^{[jj]} \sum_{i=1}^{N} \sum_{j=1}^{|\omega|} \left(k_{\sigma(j,\chi(x_i))}^{[ii]} - k_{\sigma(j,\chi(x_i))}^{[jj]}\right)^2 \left\| x_i^{[ii]} - \omega_j^{[ii]} \right\|^2 + \sum_{jj=1,jj\neq 1}^{P} \alpha_{[ii]}^{[jj]} \sum_{i=1}^{N[ii]} \sum_{j=1}^{|\omega|} \left(k_{\sigma(j,\chi(x_i))}^{[ii]} - k_{\sigma(j,\chi(x_i))}^{[jj]}\right)^2 \left\| \omega_j^{[ii]} - \omega_j^{[jj]} \right\|^2 \quad (10)$$

Where P represents the number of databases, N[ii] represents the number of individuals of the ii-th database; |w| represents the number of vectors of the referents AdSOM map. The equation (10) is minimized for each database in collaboration step.

The fundamental proposition of hybrid collaborative clustering is illustrated in algorithm 3. It can be divided into two steps: a local step and a collaborative step. In the local step, the enriched AdSOM algorithm is independently applied to each dataset. The collaboration step would combine each of the datasets with all AdSOM maps associated with other bases obtained in horizontal and vertical collaboration.

3.4 Random subspace Method

The random subspace method (RSM) is an ensemble learning that consists of several classifiers every operating in a subspace of the original attribute space introduced by Ho [16]. In the RSM, the training set is also changed as in bagging. Although, this modification is performed in the feature space. In the random subspace approach [18], classifiers are trained from random subspaces of the data feature space. In the beginning, it arbitrarily chooses a subset of the features and applies a deterministic version of the base-level method.

Each learning example X_j in the training sample set S is a p-dimensional vector $X_j = (x_{j1}, x_{j2}, ..., x_{jp})$. In the RSM, p^+ features are randomly selected from the training set S, where $p^+ < p$.

In this way, a p^+ dimensional random subspace of the original p-dimensional feature space is obtained. The resulting training set $S^{\sim} = (X_1^{\sim}, X_2^{\sim}, ..., X_n^{\sim})$ consists of p^+ -dimensional training examples $X_j^{\sim} = (x_{j1}^{\sim}, x_{j2}^{\sim}, ..., x_{jp+}^{\sim})$ $(j = 1, 2, ..., n)$. Afterwards, base-level classifiers are constructed from the random subspaces $S_i^{\sim}$, $i = 1, 2, ..., N$, and they are cooperated by a voting scheme to collect a final forecast.

The RSM algorithm is shown in algorithm 4.

Algorithm4: Random subspace method

```
Input: Training examples S, Number of subspaces N, Dimension of subspaces $p^+$

Output: Ensemble E
$E \Leftarrow 0$
for $i = 1$ to N do
    $S_i^{\sim} \Leftarrow$ SelectRandomSubspace(S, $p^+$)
    $C_i \Leftarrow$ ConstructClassifier($S_i^{\sim}$)
    $E \Leftarrow E \cup \{C_i\}$
end for
return E
```

The RSM may benefit from using both random subspaces to build classifiers and to aggregate them. Seeing that the number of training examples is comparatively small as compared with the data dimensionality, one may find solution to this problem by building classifiers in random subspaces. In this case the subspace dimensionality is lower than in the original characteristic space, while the number of training observations remains the same. When the dataset contains many redundant features, one may generate better classifiers in random subspaces than in the original feature space. The cooperative decision of such classifiers may be superior to that of a single classifier constructed on the original dataset in the complete feature space.

The RSM was initially developed for decision trees, but the ensemble learning can be applied to improve the quality of other unstable classifiers (e.g. rules, neural networks etc.). RSM performs effectively when there is a degree of redundancy in the data feature space ([29], [17]). It is noted that the performance of the RSM is designated by problem of complexity (feature efficiency, length of class boundary) [15]. While applying to decision trees, the RSM is superior to a unique classifier and may perform better than both bagging and boosting [18]. RSM is a comparatively recent method of cooperating models. Learning machines are trained on indiscriminately selected subspaces of the original input space. The results of the models are then combined generally by a simple majority vote.

4. EXPERIMENTAL RESULTS

To assess the proposed collaborative methods, the approaches are tested on the TIMIT and NTIMIT datasets. These datasets are divided into two parts: training and testing. The same partition is used to evaluate the approaches. Additional information on these datasets is provided to illustrate the principle of the proposed methods, particularly in the validation stage. As evaluation criteria, the recognition rate is used on several maps with different sizes for each SOM. A hybrid collaboration system for continuous speech recognition is implemented, based on AdSOM systems composed of three principal components: a procedure for vowel sound processing and mel-cepstrum vector generation. Then, the sound input space is comprised by 12 mel-cepstrum coefficients for every 16 ms frame. Three frames are chosen in the middle of every phoneme to create data vectors. The second constituent is a hybrid collaboration which combines both approaches of vertical and horizontal collaboration using AdSOM. The third constituent is a phoneme recognition module. The speech datasets are based on the DARPA TIMIT and NTIMIT acoustic-phonetic continuous speech corpora.

4.1 Datasets

To test and compare the performances of the proposed algorithms with respect to those obtained using ensemble methods, horizontal and vertical collaboration proposed in literature, all experiments reported below were performed using the TIMIT and NTIMIT corpora [30], [31]. These corpora are used to develop and evaluate the proposed collaborative approaches for speaker-independent continuous speech recognition. The TIMIT dataset is a reference American English speech corpus, most adopted for phone recognition and segmentation experiments on read speech.

The corpus contains 6,300 sentences (630 speakers, 10 sentences per speaker) representing 8 different US dialectal regions. The data are divided into dialect sentences (whose file name begins with SA), phonetically-compact sentences (beginning with SX) and phonetically-diverse sentences (beginning with SI). The dialect sentences are two and are produced by all the 630 speakers; they are designed to show the dialectal origin of the speaker. There are 438 male speakers and 192 female speakers. The phonetically-compact sentences (450 in all) are explicitly meant to provide a good coverage of all the American English phone pairs; each speaker reads 5 phonetically compact sentences. Finally, the phonetically-diverse sentences are defined with the aim of maximizing the variety of allophonic contexts; there are 1890 in all (3 sentences per speaker). The symbol set is divided into the following 6 categories: stops, affricates, fricatives, nasals, vowels, semivowels (and glides). Other symbols are used for pause, silence and as a marker. In the experiments, the New England dialect region (DR1), comprising 31 males and 18 females, is selected. The corpus includes 14,399 phonetic cells for training. Every phonetic cell is described by three frames picked at the middle of every phoneme to obtain data vectors. Training has been made on vowel phonemes, which are : ‘iy', ‘ih', ‘eh', ‘ey',‘ae', ‘aa', ‘aw', ‘ay', ‘ah', ‘ao', ‘oy', ‘ow', ‘uh', ‘uw', ‘ux', ‘er', ‘ax', ‘ix', ‘axr', and ‘ax-h'.

The NTIMIT corpus, was implemented by the NYNEX Science and Technology Speech Communication Group, is a telephone bandwidth version of TIMIT. It was constructed transmitting the TIMIT sentences over telephone channels, through either local or long-distance calls. This database is a precious resource for studying the effect of the telephone bandwidth noise on the performance of speech technology applications. Manual phonetic transcription is provided for both the TIMIT and NTIMIT speech corpora. In the experiments, both datasets were used, including the SA sentences.

Experiments were conducted on both TIMIT (clean speech) and NTIMIT (noisy speech) datasets under multiple noise conditions. The proposed Hybrid Collaborative Clustering (HCC) approach was compared with Random Subspace, and recognition rate (%) was used as the evaluation metric.

4.2 Data partitioning

The datasets mentioned above are unified and must be divided into subsets in order to distribute the data scenarios; vertical, horizontal, and hybrid partitioning are applied (Fig 1). In the horizontal approach, the datasets are divided into subsets such that each algorithm operates on different characteristics but on the same observations.

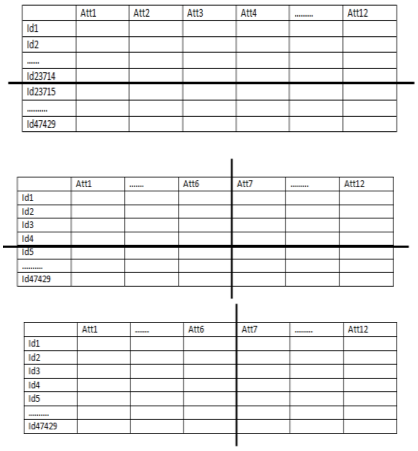


Fig 1: Vertical (First), horizontal (Third) and hybrid (Second) partitioning.

In contrast, in vertical collaboration, each method operates on the same characteristics but on different observations. Finally, in the hybrid approach, these two principles are combined, and the datasets are divided into subsets containing both different characteristics and different groups of individuals.

4.3 Interpretation of these approaches on the TIMIT and NTIMIT datasets

4.3.1 Procedure of the horizontal approach

The TIMIT and NTIMIT datasets, each of size 47429 × 12, are partitioned into four subsets to simulate a horizontal collaborative clustering scenario across the four parts. The first and the second subsets of the dataset TIMIT 2× (47429×6) contain all the pertinent features and the third and fourth subsets of the dataset NTIMIT 2× (47429×6) include noisy variables.

Maps of size 10 × 10 are used. The local step of the horizontal collaborative clustering approach is then applied to the four datasets, allowing an AdSOM to be learned from all observations in these datasets. These datasets are used to present the entire process and to enable horizontal collaboration among all datasets.

The collaboration matrix is initially set to [11; 11]. The local step of the proposed collaborative clustering algorithm is then executed on all four datasets. The two axes X and Y illustrate respectively the indices of variables and referents for these maps. Figure 2(a) and figure 2(b) present respectively the referents vectors obtained on the first SOMT1 and second SOMT2 maps after the local learning step and correspond to the maps which contain the pertinent features from the TIMIT dataset (Vowels) which are represented by the red color and have a recognition rate of 62.54% and 61.94%.

Thereafter, the local phase of the method is applied to the NTIMIT dataset, which contains noisy variables, producing the SOMNT1 and SOMNT2 maps with recognition rates of 42.05% and 45.99%, respectively (Figure 3). Subsequently, another map for the first database (SOMT1) was constructed through collaboration with the SOMNT1 map, as illustrated in Figure 4(a). A similar collaboration was carried out for the SOMNT1 map using the first map (SOMT1), resulting in the SOMNT1_T1 map, as illustrated in Figure 5(a). The recognition rates obtained are 56.12% for SOMT1_NT1 and 47.95% for SOMNT1_T1. The quality of the clustering of the SOMT1_NT1 map obtained after the collaborative SOMT1 with the classification outcome of SOMNT1 is lower than that of SOMT1. Indeed, SOMT1_NT1 used SOMNT1 information which has a lower recognition rate (42.05%) due to noise and shows that collaboration has occurred since the recognition rate decreased using the data of the NTIMIT dataset.

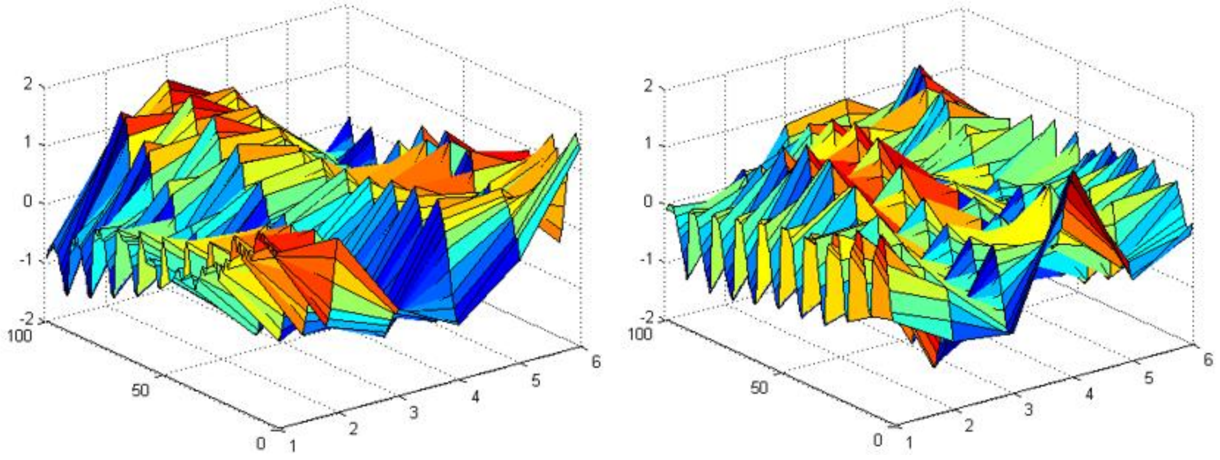
Concerning the SOMNT1_T1 map obtained after collaboration between the first noisy database and the SOMT1 map, the recognition rate increased because it utilized information from the more powerful SOMT1 map, achieving 62.54%.

The collaboration phase between the SOMT2 map and the SOMNT2 map is applied. As a result, the two maps shown in Figures 4(b) and 5(b) are obtained. After the collaboration of the second dataset (TIMIT) with irrelevant SOMNT2 map, the recognition rate decreased to 60.50% as the SOMNT2 map used information of a noisy map with a very low recognition rate (45.99%). However, by applying the collaborative phase in the opposite direction, the recognition rate of the SOMNT2_T2 map increased to 54.95 % through collaboration with relevant SOMT2 map (61.94% recognition rate). To increase the influence of the relevant map during the process of collaboration, the collaborative setting should be enhanced to the board having a greater recognition rate and decreased for the non-relevant map in order to weaken his contribution during the collaboration. For example, the recognition rate is improved from 42.05% to 56.12% by applying collaboration between the relevant SOMT1 map and the non-relevant SOMNT1 map, using collaboration indices [8; 0.2]. Horizontal collaboration is a challenging procedure because, within an unsupervised classification framework, identifying the most relevant maps is difficult. Consequently, collaboration must be performed in both directions, followed by a comparison and validation of the results.

Table 1 presents the recognition rates of the different Vowels of the datasets TIMIT and NTIMIT before and after the collaboration. The recognition rates are improved (increased) after collaboration with a more relevant map (a richer classification).

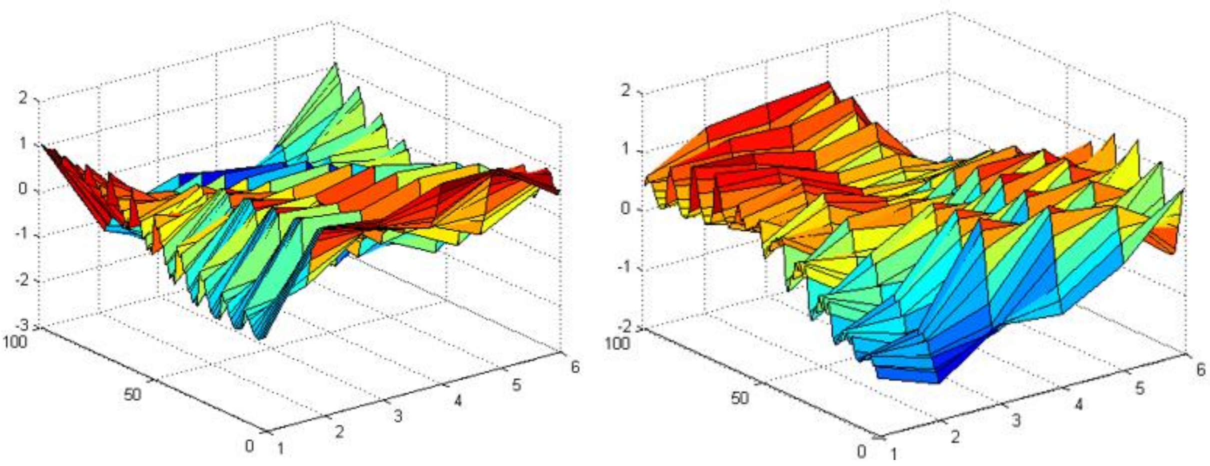
4.3.2 Procedure of the vertical approach

For vertical collaboration on the TIMIT and NTIMIT datasets, the data were partitioned into four sub-datasets. Unlike horizontal collaboration, this partitioning was performed randomly across individuals. Consequently, four sub-datasets of size (23714 × 12) were obtained for each dataset. Furthermore, an identity collaboration matrix was established. Maps of dimension 10 × 10 are chosen. The results obtained in this experiment of vertical collaboration using the TIMIT and NTIMIT datasets are summarized in Table 2. From Table 2, it can be observed that in most cases, for maps obtained after collaboration with more relevant maps, the recognition rate increases, as is the case for SOMNT1_T1 and SOMNT2_T2. Since all the four bases are described in the same area (the same 12 variables), recognition rate maps before and after the contribution is large enough. In addition, the collaboration confidence parameters are comparative because all the maps are uniform. The recognition rate of the maps before and after the collaboration is superior to horizontal collaboration because the four databases are described in the same variables. These results illustrate that the recognition rate of maps depends effectively on the pertinence of the collaborative map and on the confidence on this map. This conclusion is consistent with the intuitive understanding of the concept and the observed results of the collaboration. In most cases, the recognition rate of the map improves following collaboration.



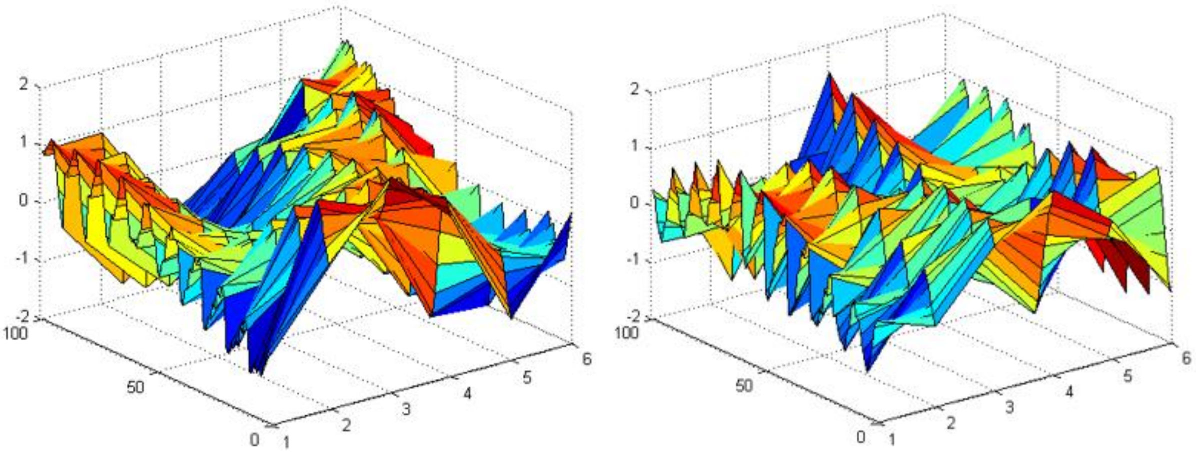
(a) SOMT1 (TIMIT1) **(b) SOMT2 (TIMIT2)**

Fig 2: Horizontal collaboration: local phase for dataset TIMIT.



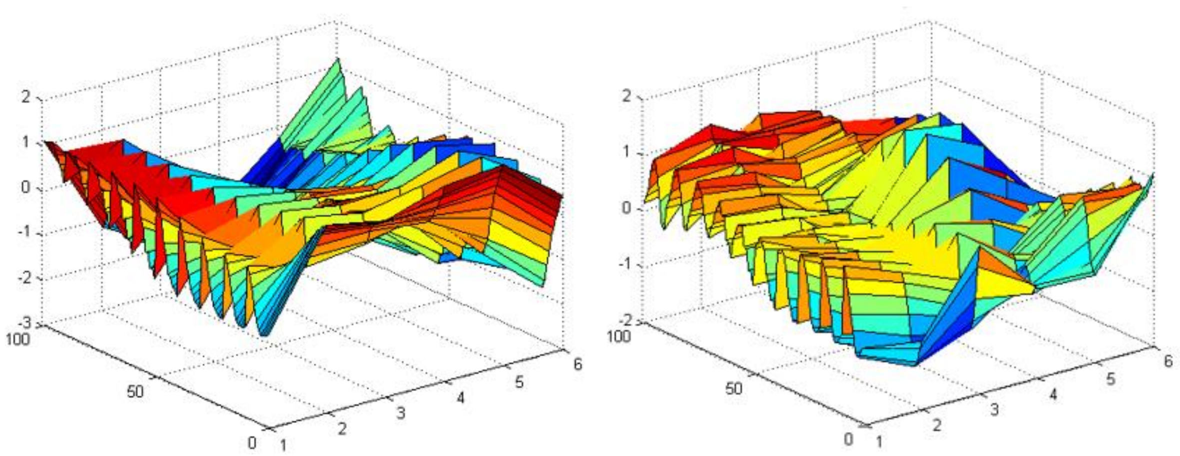
(a) SOMNT1 (NTIMIT1) **(b) SOMNT2 (NTIMIT2)**

Fig 3: Horizontal collaboration: local phase for dataset NTIMIT.



(a) SOMT1_NT1(TIMIT1-NTIMIT1) **(b) SOMT2_NT2 (TIMIT2-NTIMIT2)**

Fig 4: Horizontal collaboration: collaboration phase for datasets TIMIT and NTIMIT.



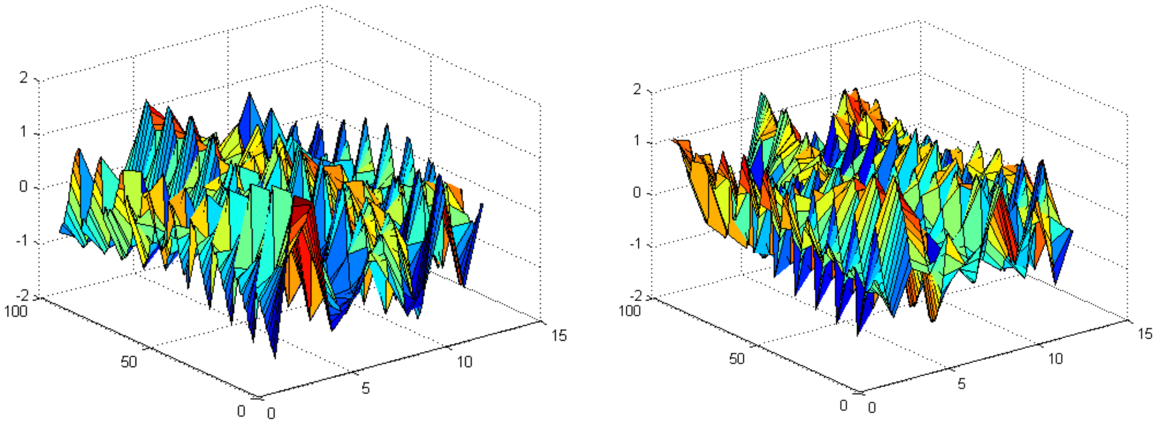
(a) SOMNT1_T1(NTIMIT1-TIMIT1) **(b) SOMNT2_T2 (NTIMIT2-TIMIT2)**

Fig 5: Horizontal collaboration: collaboration phase for datasets NTIMIT and TIMIT.

Table 1. Recognition rates of horizontal collaboration (datasets TIMIT and NTIMIT).

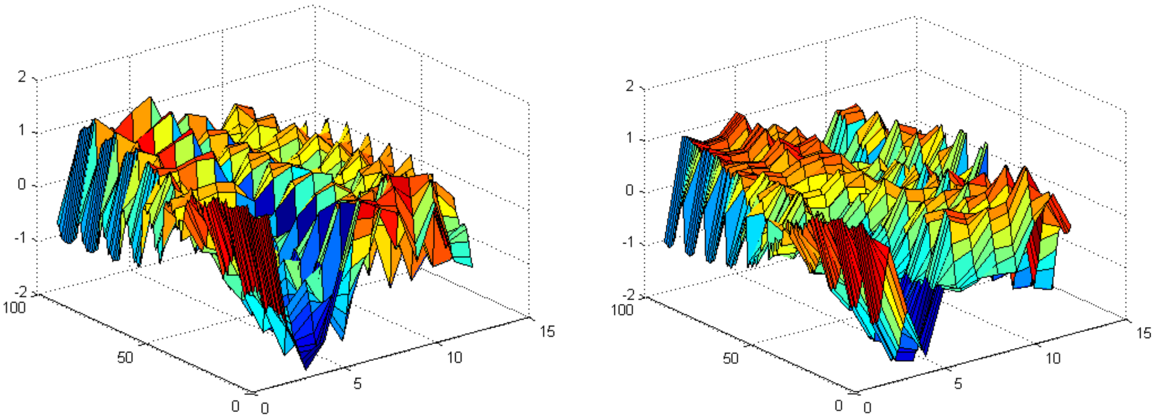
	SOM T 1	SOM T2	SOM NT1	SOM NT2	SOM T1_NT1	SOM T2_NT2	SOM NT1_T1	SOM NT2_T2
iy	85,36	79,49	64,45	39,95	**76,88**	**77,89**	**71,02**	**68,39**
ih	69,65	67,22	59,05	51,7	**65,74**	**65,77**	**62,22**	**62,21**
eh	80,66	98,57	68,63	79,12	**78,54**	**95,33**	**73,45**	**91,74**
ey	95,37	79,35	52,72	65,88	**89,66**	**78,62**	**62,11**	**77,12**
ae	78,75	73,86	73,19	52,32	**77,21**	**70,98**	**74,87**	**66,33**
aa	57,94	58,24	47,17	42,88	**53,81**	**55,71**	**48,55**	**49,56**
aw	39,77	29,83	19,44	27,31	**32,77**	**29,11**	**25,77**	**28,31**
ah	77,33	75,26	52,52	58,77	**67,03**	**72,14**	**59,87**	**61,78**
ay	38,16	36,53	29,87	31,66	**35,21**	**35,44**	**32,44**	**34,11**
ao	66,19	62,81	20,80	51,77	**49,58**	**60,78**	**30,74**	**55,45**
oy	70,3	84,50	61,77	42,12	**68,66**	**80,52**	**64,22**	**71,02**
ow	69,95	51,93	34,07	42,32	**55,22**	**49,61**	**45,71**	**45,66**
uh	72,66	83,87	40,98	50,37	**62,07**	**80,01**	**49,5**	**66,84**
uw	49,02	45,17	32,50	44,65	**41,76**	**45,04**	**36,44**	**44,86**
ux	59,78	61,70	50	45,23	**57,74**	**58,44**	**54**	**51,33**
er	57,8	64,27	26,26	51,89	**49,89**	**62,14**	**38,21**	**55,22**
ax	86,6	77,98	25,15	49,37	**69,33**	**75,22**	**44,17**	**68,65**
ix	57,28	67,12	51,22	57,86	**56,2**	**76,54**	**53,66**	**62,11**
axr	38,39	41,21	31,33	34,75	**35,1**	**40,75**	**32,09**	**38,41**
axh	0	0	0	0	**0**	**0**	**0**	**0**

Auto-coherence rates%	62.54	61.94	42.05	45.99	**56.12**	**60.50**	**47.95**	**54.95**
Generalisation rates%	59.33	58.78	40.45	43.77	**55.75**	**57.66**	**46.32**	**52.88**



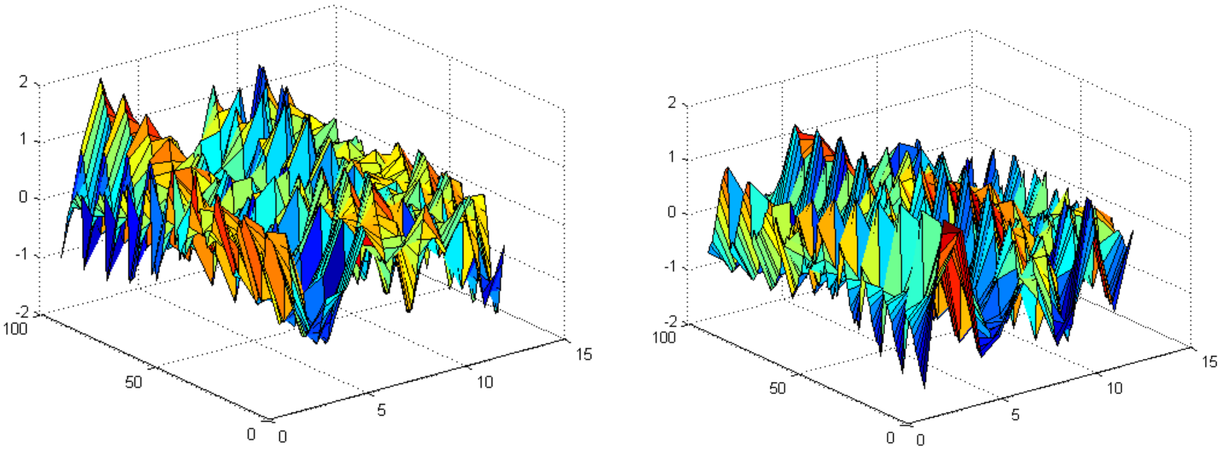
(a) SOMT1(TIMIT1) **(b) SOMT2 (TIMIT2)**

Fig 6: Vertical collaboration: local phase for dataset TIMIT.



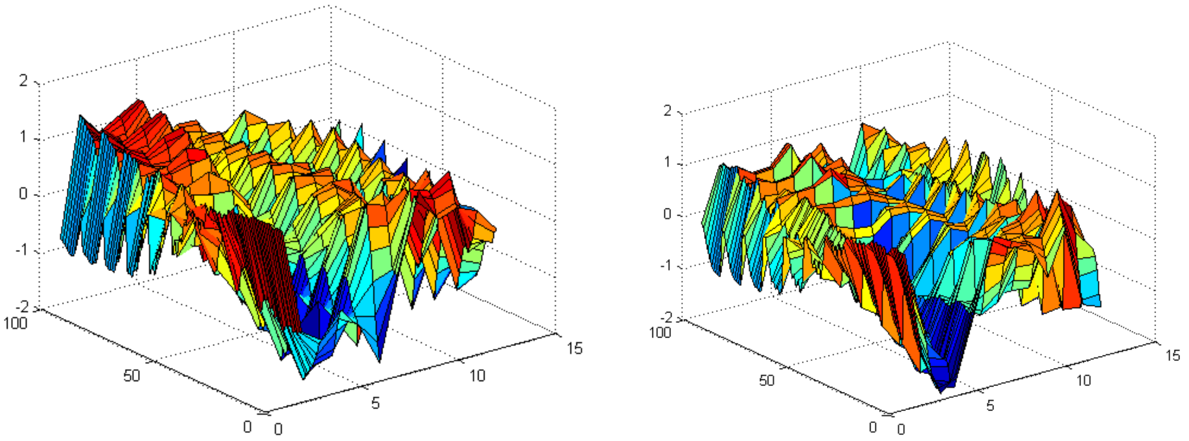
(a) SOMNT1(NTIMIT1) **(b) SOMNT2 (NTIMIT2)**

Fig 7: Vertical collaboration: local phase for dataset NTIMIT.



(a) SOMT1_NT1(TIMIT1-NTIMIT1) **(b) SOMT2_NT2 (TIMIT2-NTIMIT2)**

Fig 8: Vertical collaboration: collaboration phase for datasets TIMIT and NTIMIT.

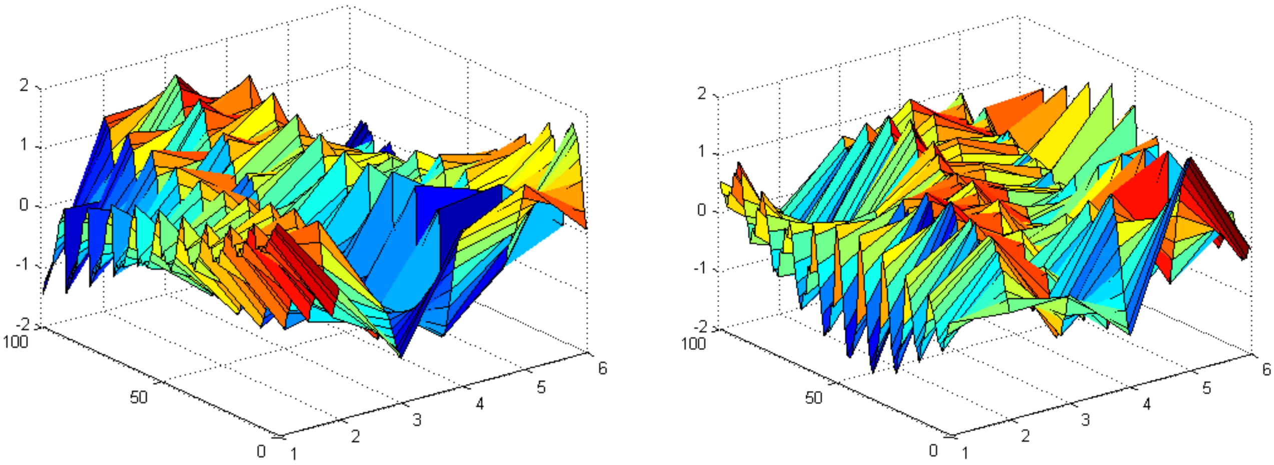


(a) SOMNT1_T1(NTIMIT1-TIMIT1) **(b) SOMNT2_T2(NTIMIT2-TIMIT2)**

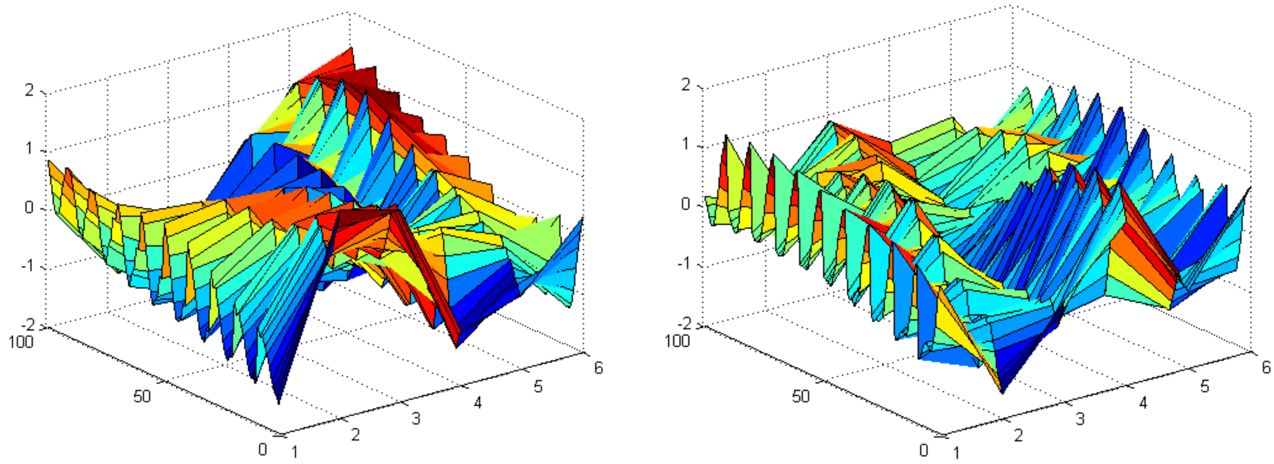
Fig 9: Vertical collaboration: collaboration phase for datasets NTIMIT and TIMIT.

Table 2. Recognition rates of vertical collaboration (datasets TIMIT and NTIMIT)

	SOM T1	**SOM T2**	**SOM NT1**	**SOM NT2**	**SOM T1_NT1**	**SOM T2_NT2**	**SOM NT1_T1**	**SOM NT2_T2**
iy	83.10	86.03	57.76	70.8	**81.45**	**84.76**	**67.52**	**74.25**
ih	87.38	91.80	80.22	76.21	**85.32**	**83.66**	**81.99**	**81.22**
eh	95.11	83.23	64.30	79.25	**98.56**	**81.30**	**79.82**	**81.01**
ey	97.20	93.13	89.50	34.22	**94.94**	**76.27**	**91.30**	**55.10**
ae	73.09	75.18	59.38	65.18	**61.24**	**74.84**	**63.41**	**73.89**
aa	77.21	80.58	38.7	57.34	**64.30**	**75.80**	**45.23**	**75.20**
aw	92.87	86.88	28.39	20.32	**53.49**	**58.30**	**46.35**	**33.54**
ah	84.18	76.11	28.33	36.29	**60.77**	**46.74**	**32.91**	**48.98**
ay	89.73	80.51	21.83	41.16	**51.05**	**71.81**	**38.76**	**64.68**
ao	63.83	92.16	59.23	43.47	**59.49**	**70.55**	**62.72**	**55.33**
oy	44.22	53.19	27.23	15.90	**37.40**	**44.59**	**35.33**	**25.77**
ow	74.38	92.87	50.70	72.32	**56.08**	**89.83**	**60.67**	**82.84**
uh	71.11	64.84	49.52	53.13	**60.76**	**59.04**	**55.60**	**58.64**
uw	56.25	62.51	48.15	54.16	**53.99**	**60.07**	**53.21**	**56.01**
ux	72.05	85.89	37.93	66.3	**50.40**	**76.54**	**52.44**	**75.25**
er	80.15	84.43	62.41	75.09	**74.88**	**82.64**	**73.11**	**79.52**
ax	75.57	73.25	70.61	25.84	**71.88**	**55.67**	**72.45**	**42.01**
ix	73.54	92.46	29.20	25.95	**57.66**	**67.33**	**38.75**	**35.35**
axr	54.89	23.89	18.72	10.16	**46.30**	**20.46**	**27.99**	**13.82**
axh	0	0	0	0	**0**	**0**	**0**	**0**
Auto-coherence rates%	72.29	73.94	46.10	46.15	**60.99**	**64.01**	**53.97**	**55.62**
Generalisation rates%	70.14	72.58	45.45	45.66	**60.45**	**62.79**	**52.87**	**55.12**

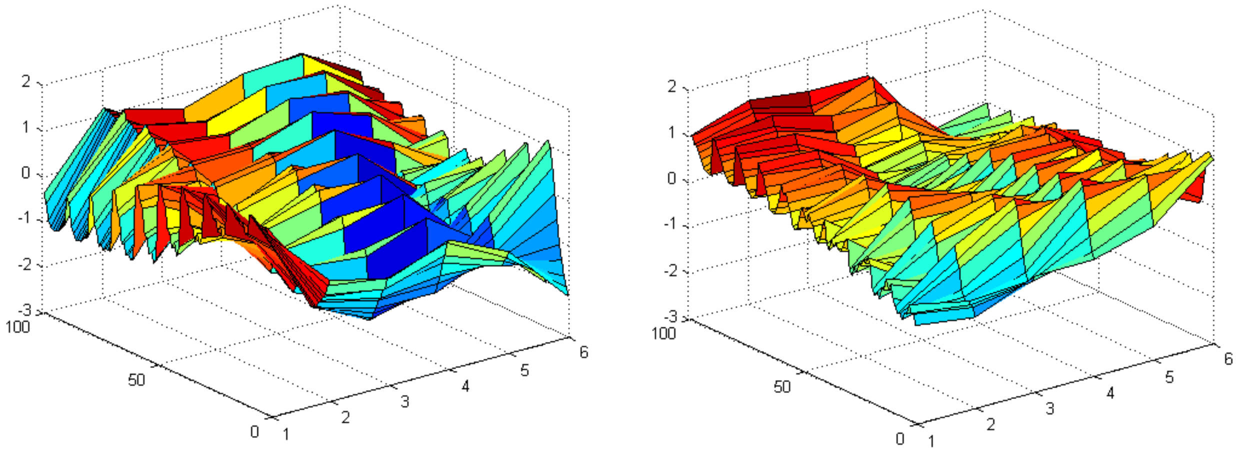


(a) SOMT1 (TIMIT1) **(b) SOMT2 (TIMIT2)**

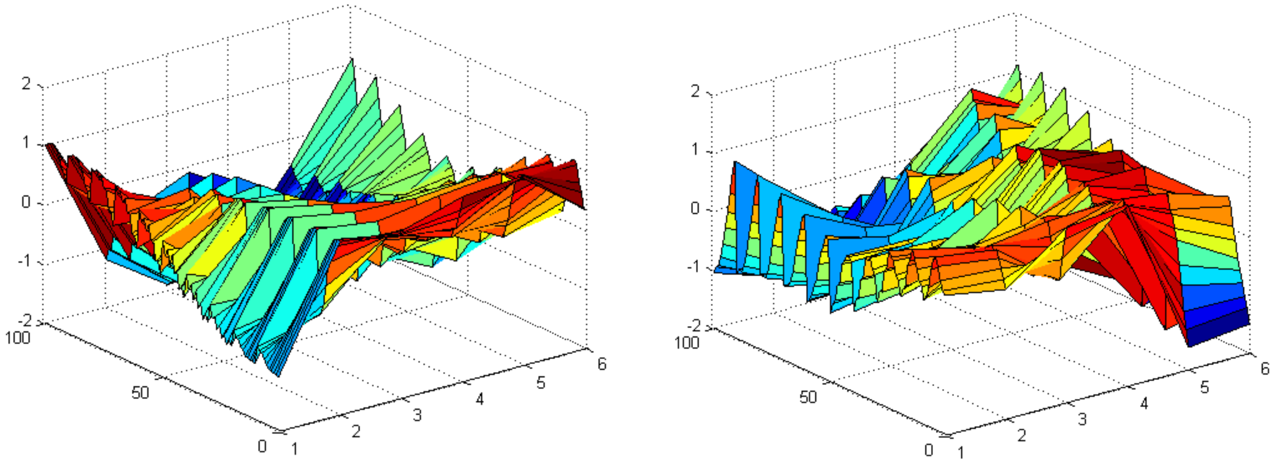


(c) SOMT3 (TIMIT3) **(d) SOMT4 (TIMIT4)**

Fig 10: Hybrid collaboration: local phase for dataset TIMIT.

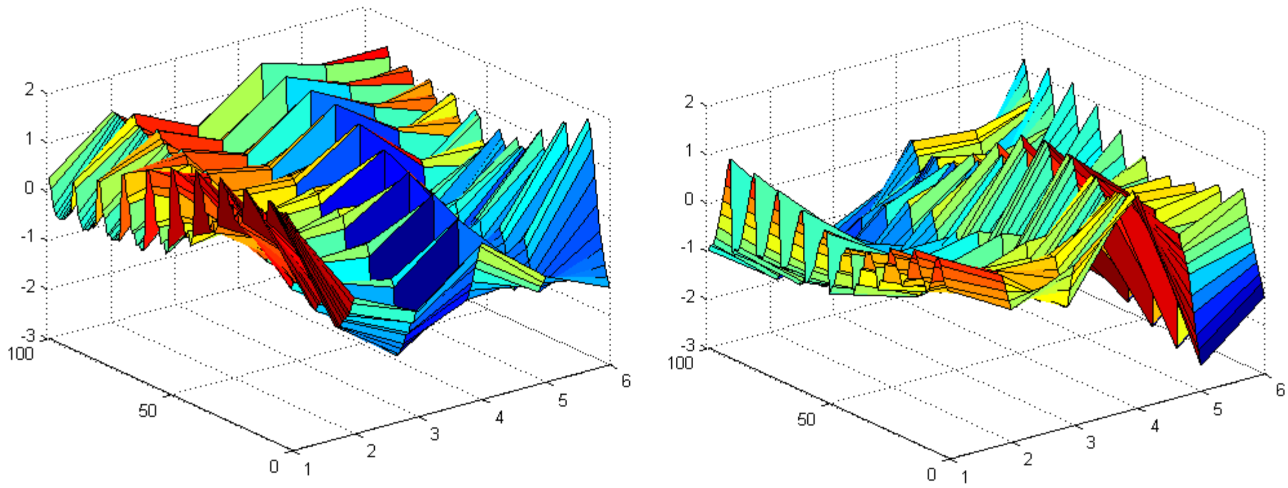


(a) SOMNT1 (NTIMIT1) **(b) SOMNT2 (NTIMIT2)**



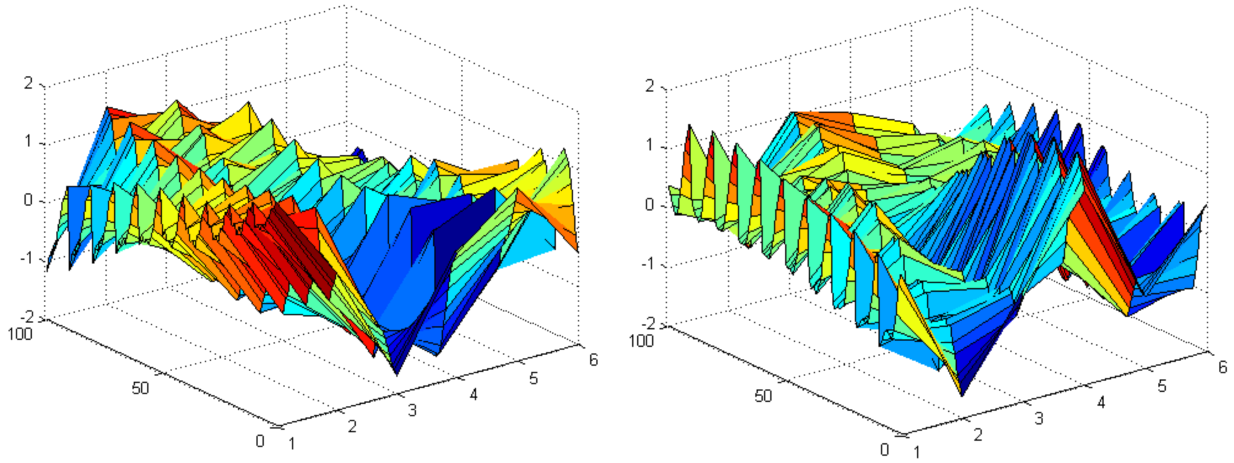
(c) SOMNT3 (NTIMIT3) **(d) SOMNT4 (NTIMIT4)**

Fig 11: Hybrid collaboration: local phase for dataset NTIMI



(a) SOMT1_NT1 (TIMIT1-NTIMIT1) **(b) SOMT4_NT4 (TIMIT4-NTIMIT4)**

Fig 12: Hybrid collaboration: collaboration phase for datasets TIMIT and NTIMIT.



(a) SOMNT1_T1 (NTIMIT1-TIMIT1) **(b) SOMNT4_T4 (NTIMIT4-TIMIT4)**

Fig 13: Hybrid collaboration: collaboration phase for datasets NTIMIT and TIMIT.

Table 3. Recognition rates of hybrid collaboration (datasets TIMIT and NTIMIT)

	TIMIT				**NTIMIT**				**Hybrid Collaboration**			
	SOM T1	**SOM T2**	**SOM T3**	**SOM T4**	**SOM NT1**	**SOM NT2**	**SOM NT3**	**SOM NT4**	**SOM T1_NT1**	**SOM T4_NT4**	**SOM NT1_T1**	**SOM NT4_T4**
iy	76,57	79,71	77,05	72,91	50,80	56,42	53,26	54,73	**70,87**	**71.09**	**64,28**	**63,78**
ih	85,44	97,38	95,84	91,79	45,24	54,70	58,60	52,91	**75,32**	**78,24**	**59,30**	**65,32**
eh	95,27	98,29	83,43	78,66	62,67	63,53	57,57	64,36	**87,12**	**72,77**	**70,89**	**71,06**
ey	96,42	89,39	88,66	97,75	57,19	50,77	59,80	64,49	**82,93**	**86,36**	**68,24**	**75,36**
ae	98,90	78,36	83,93	79,56	58,63	60,09	63,94	59,23	**85,69**	**76,21**	**71,22**	**69,87**
aa	97,09	78,98	81,37	94,37	47,83	40,61	54,37	48,66	**79,36**	**82,45**	**65,74**	**62,33**
aw	42,89	43,49	40,99	45,77	39,87	36,11	31,87	35,31	**42,25**	**42,56**	**40,95**	**41,45**
ah	77,35	75,51	77,86	76,82	67,39	50,38	56,32	49,39	**75,82**	**70,23**	**72,03**	**67,75**
ay	47,84	48,02	45,74	45,38	38,88	35,64	30,59	32,47	**45,32**	**42,21**	**43,66**	**41,21**
ao	86,65	82,59	80,31	79,87	50,59	63,35	69,15	58,75	**72,79**	**75,36**	**60,45**	**65,43**
oy	75,34	66,69	69,95	74,28	38,26	40,57	54,47	51,62	**66,36**	**70,45**	**47,25**	**60,58**
ow	85,63	91,62	89,64	79,74	54,57	45,71	46,73	53,47	**76,45**	**73,55**	**65,33**	**63,14**
uh	79,28	86,33	79,29	80,19	45,77	42,14	48,54	45,88	**63,24**	**71,78**	**55,74**	**57,68**
uw	60,92	55,58	72,68	73,88	51,46	50,74	52,42	54,13	**56,78**	**67,25**	**56,54**	**62,31**
ux	59,78	65,16	63,75	65,74	47,58	41,26	53,50	41,77	**55,89**	**55,37**	**53,78**	**55,65**
er	86,39	82,29	89,19	88,28	47,74	43,91	38,21	40,45	**71,21**	**68,98**	**62,76**	**60,55**
ax	86,64	81,46	86,31	87,39	64,74	60,28	46,80	57,49	**78,36**	**71,68**	**75,21**	**64,25**
ix	96,93	85,85	92,13	81,46	36,19	53,59	58,86	58,18	**70,30**	**70,03**	**56,45**	**61,54**
axr	49,17	56,58	55,39	56,56	42,61	43,41	37,44	39,59	**46,74**	**48,25**	**45,33**	**46,36**
axh	0	0	0	0	0	0	0	0	**0**	**0**	**0**	**0**
Auto-coherence rates %	74.22	72.16	72.67	72.52	47.40	46.66	48.62	48.14	**65,14**	**61,18**	**56,75**	**57,78**
Generali-sation rates	73.85	71.55	72.10	71.69	45.82	45.54	47.85	47.62	**64.38**	**60.78**	**56.32**	**56.98**

Table 4. Recognition rates of Random Subspace (datasets TIMIT and NTIMIT)

	Random subspace (TIMIT)	**Random subspace (NTIMIT)**
iy	71,17	51,88
ih	55,00	47,97
eh	60,42	42,91
ey	60,83	53,67
ae	59,81	47,07
aa	56,82	49,88
aw	59,87	53,79
ah	70,91	64,19
ay	53,78	47,07
ao	69,57	56,79
oy	63,51	52,28
ow	58,96	47,61
uh	66,52	53,74
uw	64,39	52,28
ux	73,78	49,79
er	57,43	43,12
ax	64,08	56,19
ix	70,22	43,58
axr	58,04	48,97
axh	0	0
Rate %	59.75	48.13

4.3.3 Procedure of the hybrid approach

This section presents the proposed contribution, which combines both topological collaborative approaches using Self-Organizing Maps in horizontal and vertical collaboration, as described in the previous section. In the hybrid collaboration section, the TIMIT and NTIMIT databases, each of size (47429 × 12), were partitioned into a grid of eight subsets to implement a hybrid collaborative clustering framework among these subsets.

The first four parts of the dataset TIMIT 4× (23714×6) correspond to all the relevant features and from the fifth part to the eighth part of the dataset NTIMIT 4× (23714 ×6) contain noisy variables. The map size is set to 10 × 10. Subsequently, the local phase of the hybrid collaborative clustering is applied to all databases, enabling a SOM to be learned from all objects in these datasets. These datasets are then used to present the entire process, enabling hybrid collaboration among all databases. Initially, the collaboration coefficients were set to [11; 11]. The local step of the collaborative clustering method is applied to all datasets following the local step of the new clustering method. Figure 10 presents respectively the referents vectors obtained on the four maps after the local learning step and correspond to the maps which contain the pertinent features from the TIMIT dataset (Vowels) which are represented by the red color and have a recognition rate of 74.22%, 72.16%, 72.67% and 72.52% respectively. Afterwards, the local phase of the approach is applied to the four NTIMIT subsets containing noisy features, producing non-pertinent maps with recognition rates of 42.61%, 43.41%, 37.44%, and 39.59%, respectively. Subsequently, another map for the first dataset (TIMIT) was constructed through collaboration with the first map (NTIMIT), as illustrated in Figure 12(a). A similar collaboration was carried out for the first dataset (NTIMIT) using the first map (TIMIT), resulting in the SOMNT1_T1 map, as shown in Figure 13(a). The recognition rates of maps are: 65.14% for the map SOMT1_NT1 and 56.75 % for the map SOMNT1_T1. The quality of the clustering of the SOMT1_NT1 map obtained after the collaborative SOMT1 with the clustering outcome of SOMNT1 is lower than that outcome of SOMT1. Indeed, SOMT1_NT1 used SOMNT1 information which has a lower recognition rate (42.61%) due to noise and shows that collaboration has occurred since the recognition rate decreased using the data of the 1st dataset NTIMIT.

The two maps shown in Figures 10(d) and 11(d) are obtained following the collaboration phase between the SOMT4 map and the SOMNT4 map. After the collaboration of the fourth dataset (TIMIT) with irrelevant map SOMNT4 (NTIMIT), the recognition rate decreased to 61.18% as the SOMNT4 map used information of a noisy map with a very low recognition rate (39.59%).

Yet, by applying the collaborative phase in the opposite direction, the recognition rate of the SOMNT4_T4 map increased to 57,78 % through collaboration with relevant map SOMT4 (72.52 % recognition rate). It is clear that hybrid collaboration yields higher results than both horizontal and vertical collaboration, owing to the combination of these two collaborative approaches. Hybrid collaboration provides best recognition rates in comparison with other collaboration approaches using AdSOM. Table 3 illustrates the recognition rates of the different Vowels of the datasets TIMIT and NTIMIT before and after the hybrid collaboration. The recognition rates are improved (increased) after collaboration with a more relevant map.

To provide a comprehensive evaluation, the proposed HCC approach was tested on both TIMIT and NTIMIT datasets under various noise conditions. Recognition rates were compared with the Random Subspace method to assess performance across clean and noisy environments.

4.3.4 Procedure of the Random Subspace

In this section, experiments are performed to discuss and analyze the behavior of the basic Random Subspace Method (RSM) compared to the horizontal collaboration approach based on SOM and other combining methods with the datasets TIMIT and NTIMIT. RSM [9] is a relatively recent method of combining models. Learning machines are trained on randomly chosen subspaces of the original input space (i.e. the training set is sampled in

the feature space). The results of the models are then cooperated, usually by a simple majority vote. RSM use SOM as the base classifier. In this work, a comparison is conducted among three ensemble learning algorithms: the hybrid approach, the horizontal collaboration approach, and RSM, as these methods operate in the feature space. The recognition rate is used for the experimental evaluation of these approaches. It is clearly seen that the TIMIT labeling reflects the hybrid and the horizontal collaboration better than the random subspace method. Whereas with the corpus NTIMIT, the random subspace method gives more effective results than horizontal collaborative approach but lower than hybrid collaborative. Indeed, RSM has a higher recognition rate (48.13%) in the noisy dataset NTIMIT compared to horizontal collaboration that has a recognition rate (47.95%). However, for dataset TIMIT, RSM has a lower recognition rate (59.75%) than the rate of hybrid collaboration (65.14%). Table 4 shows the different results obtained by applying RSM on the both datasets.

RSM proves more robust than horizontal collaboration in noisy settings. For this reason, a comparison is conducted between horizontal collaboration, RSM, and other well-known ensemble methods on the TIMIT and NTIMIT datasets. For the experiments, representative algorithms of well-known machine learning techniques, such as Support Vector Machines, were also included in the evaluation.

While our experiments were conducted on TIMIT and NTIMIT, the proposed Hybrid Collaborative Clustering approach is generalizable and could be evaluated on larger or more diverse datasets such as **LibriSpeech, VCTK, CHiME, Aurora, or WSJ** for further validation.

5. CONCLUSION

Speech phonetic collaboration is a challenging task that finds application in speech recognition, coding and synthesis. Collaborative clustering is helpful to reach interaction between different datasets for the object of revealing underlying structures and regularities. Hence, it can be considered a process of consensus construction where a common structure across all datasets is revealed. The proposed horizontal and vertical clustering methods achieve an active form of collaboration. The focus is on data security and confidentiality: the level of granularity at which communication takes place provides a practical means of preserving these properties. The principle of collaborative clustering could be employed in the design of intelligent agents operating within limited domains, allowing them to benefit from different forms of interaction and collaboration. This process must occur in an active mode, which is exactly what collaborative clustering ensures. Random subspace, Vertical, horizontal, and hybrid collaborative clustering are fundamental mechanisms of interaction at some granular level. Various collaborative activities furnish insight into the structure of data and help to identify patterns that require more attention or could be labeled as outliers. This work proposes an original collaboration algorithm that performs phonetic collaboration between the TIMIT (pertinent) and NTIMIT (non-pertinent) datasets using horizontal and vertical approachesThe combination of these two approaches is also proposed to design a new hybrid collaborative method based on AdSOM for phoneme recognition.

In addition, random subspaces are applied to these two datasets, and a comparison is made with the hybrid and horizontal collaboration approaches. The main results are as follows:

- In vertical collaboration, as all databases are described in the same variables, the recognition rate of the maps before and after the collaboration is greater compared to the horizontal collaboration.
- In hybrid collaboration, where both vertical and horizontal collaborative clustering are combined, the recognition rates of the maps before and after collaboration are higher than those obtained using vertical and horizontal collaboration, as well as RSM.
- The random subspace methods are much more robust than the horizontal collaboration in noisy settings (NTIMIT).
- Horizontal collaboration performs better than the random subspace method on the TIMIT corpus. The collaborative clustering models are designed under the assumption that each dataset has the same number of clusters. Although this assumption is not particularly restrictive, it may be relaxed in some cases, especially when learning granular mappings where the assumption made so far may not be fully justified.

Collaborative clustering can be related to conditional clustering, which does not support active interaction but allows different levels of granularity and different numbers of clusters in each dataset.

Several perspectives can be considered for this work as: to apply a generative model, which is the Generative Topographic Mappings (GTM); to merge all the clustering results obtained after collaboration and to construct a consensus across all data views.

6. REFERENCES

[1] Kaur, P. and Sharma, R., 2024. Enhancing Speech Recognition Accuracy with Kernel Methods. Journal of Machine Learning Research.

[2] A. Abas, A. Mahmood, T. Rashid, H. Veisi, Advanced Clustering Techniques for Speech Signal Enhancement: A Review and Meta- analysis of Fuzzy C-Means, K-Means, and Kernel Fuzzy C-Means Methods, https://arxiv.org/pdf/2409.19448.

[3] Chen, X., et al., 2022. Hybrid Clustering Models for Speech Signal Processing. ACM Transactions on Speech and Language Processing.

[4] Ahmed, S. and Hassan, M., 2023. Clustering Methods for Improved Speaker Identification. IEEE Access.

[5] Li, X. and Wu, Y., 2024. Dynamic Clustering Algorithms for Real-Time Speech Processing. Journal of Artificial Intelligence Research.

[6] A. Filali, C. Jlassi, N. Arous,"A Hybrid Collaborative Clustering using Self-Organizing Map", 14th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA 2017), Hammamet-Tunisia

[7] W.Pedrycz , "collaborative fuzzy clustering, pattern recognition", lett. 23, p. 675-686, 2002.

R. Falcon, G.Jeon, R.Bello and J. Jeong, " Learning collaboration links in a collaborative fuzzy clustering environment", In Proceeding of the artificial intelligence 6th Mexican international conference on Advances in artificial intelligence, 2007.

[8] N. Grozavu, M.Ghassany, Y.Bennani, "Learning confidence exchange in Collaborative Clustering," IJCNN, 2011.

[9] J. Kennedy and R. Eberhart, " Particle swarm optimization", In Neural Networks. Proceedings., IEEE International Conference, 1995.

[10] J.Sublime, N.Grozavu,Y.Bennani and A.Cornuéjols, "Collaborative clustering with heterogeneous algorithms," IJCNN , 2015.

[11] L. Zijing and M. Barahona "Graph-based data clustering via multiscale community detection". Applied Network Science, Springer, 2020.

[12] B. Depaire, R. Falcon, K. Vanhoof and G. Wets, " PSO driven collaborative clustering: A clustering algorithm for ubiquitous environments". Intell. Data Anal, 2011.

[13] F.Yu, J.Tang, F.Wu, and Q. Sun, " Auto-weighted Horizontal Collaboration Fuzzy Clustering".

[14] M.Sewell," Ensemble Learning" , Department of Computer Science University College London, 2008.

[15] T.K.Ho, "The random subspace method for constructing decision forests", IEEE Transactions on Pattern Analysis and Machine

[16] P.Panov and S.Dˇzeroski, "Combining Bagging and Random Subspaces to Create Better Ensembles".

[17] S. Kotsiantis and D. Kanellopoulos, "COMBINING BAGGING, BOOSTING AND RANDOM SUBSPACE ENSEMBLES FOR REGRESSION PROBLEMS", ICIC International N 1349-4198 , 2012.

[18] T. Kohonen," Self-organizing maps", Third Edition, Springer, 2001.

[19] J. Kangas, T. Kohonen and J.Laraksonen, "Variants of self-organizing maps", IEEE Transaction on Neural Networks, 1990.

[20] D. Miljković, "Brief review of self-organizing maps", 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), DOI: 10.23919/MIPRO40886. 2017.

[21] F. Zehraoui and Y. Bennani, "New self-organizing maps for multivariate sequences procing", 439–456.879, 2008.

[22] Y. Chenglong, R. Ghanadanb and J. Ding, « Meta Clustering for CollaborativeLearning », Journal of Computational and Graphical Statistics Volume 32, 2023 - Issue 3.

[23] W. Pedrycz and P. rai, "Collaborative clustering with the use of fuzzy c-means and its quantification", fuzzy sets and systems, doi: 10.1016/j.fss.2007.12.030, 2008.

[24] M. Avriel, "Nonlinear Programming: Analysis and Methods", Prentice-Hall, Englewood Cli_s, NJ, 1976.

[25] Kong, G., Ma, Y., Xing, Z., Xin, X. , "Multi-view clustering algorithm based on feature learning and structure learning Neurocomputing, 2024.

[26] A. Cornuéjols, C. Wemmert, P. Gançarski and Y.Bennani, “Collaborative clustering: Why, when, what and how”, Information Fusion Elsevier, 2018. https://doi.org/10.1016/j.inffus.2017.04.008.

[27] A. FILALI, C.JLASSI and N.AROUS, “SOM Variants for Topological Horizontal Collaboration”, International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 10.1109/ATSIP.2016.7523117, 21-23 March 2016.

[28] S. Kotsiantis, “Random Subspace Local Projections”, Artif Intell Rev 35:223–240 DOI 10.1007/s10462-010-9192-8, 2011.

[29] J.Jankowski and al.,“NTIMIT: A Phonetically Balanced,Continuous Speech, Telephone Bandwidth Speech Database”, ICASSP ,1990.

[30] J.Garofolo, S.Lamel, L.Fisher, W. M. Fiscus, J. G. Pallett and D. S. Dahlgren, “ The DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus”, CDROM NTIS order number PB91-100354, 1992.

[31] A.FILALI, C.JLASSI and N.AROUS, “vertical collaborative clustering using Enriched Self-Organiziong Maps”, 4th International Conference on Automation, Control Engineering and Computer Science (ACECS-2016), Hammamet-Tunisia.