

Document downloaded from:

<http://hdl.handle.net/10251/180932>

This paper must be cited as:

Safont, G.; Salazar Afanador, A.; Vergara Domínguez, L. (2019). New applications of late fusion methods for EEG signal processing. IEEE. 617-621.
<https://doi.org/10.1109/CSCI49370.2019.00116>



The final publication is available at

<https://doi.org/10.1109/CSCI49370.2019.00116>

Copyright IEEE

Additional Information

New applications of late fusion methods for EEG signal processing

Gonzalo Safont*, Addisson Salazar, Luis Vergara

Institute of Telecommunications and Multimedia Applications, Universitat Politècnica de València, Valencia, Spain
gonsaar@upvnet.upv.es; asalazar@dcom.upv.es; lvergara@dcom.upv.es

Abstract—Decision fusion consists in the combination of the outputs of multiple classifiers into a common decision that is more precise or stable. In most cases, however, only classical fusion techniques are considered. This work compares the performance of several state-of-the-art fusion methods on new applications of automatic stage classification of several neuropsychological tests. The tests were staged into three classes: stimulus display, retention interval, and subject response. The considered late fusion methods were: alpha integration; copulas; Dempster-Shafer combination; independent component analysis mixture models; and behavior knowledge space. Late fusion was able to improve the performance for the task, with alpha integration yielding the most stable result.

Keywords—alpha integration, classification, late fusion, EEG, neuropsychological tests

Type of submission— Full/Regular Research Paper

Symposium—Signal & Image Processing, Computer Vision & Pattern Recognition (CSCI-ISPC)

I. INTRODUCTION

Decision fusion is a kind of data fusion that combines the decisions of multiple classifiers into a common decision about the analyzed data [1,2]. This is commonly known as “late fusion” as opposed to “early fusion”, where features from multiple sensors or modalities are combined before classification. There are several advantages to the fusion of multiple classifiers, such as improved classification performance, increased confidence, or enhanced reliability [3].

There have been several works on decision fusion on biomedical applications, such as: colonic polyp detection in CT colonography [4]; identification of auditory and visual perception processes [5]; and automatic sleep staging [6]. Most of these applications have considered simple decision fusion techniques (e.g., majority voting and score averaging). While these techniques are simple and robust, performance could potentially be improved with the usage of more sophisticated and powerful methods.

In this paper, we compare the relative performance of several state-of-the-art methods of late fusion for automatic staging of electroencephalographic signals from patients performing several learning and memory tasks. More specifically, we consider alpha integration [7-10], copulas [11], independent component analysis mixture models [12], Dempster-Shafer combination [13], and behavior knowledge space [14]. We show the capabilities of late fusion to improve the staging performance by fusing several classifiers and discuss their relative merits and strengths.

This work was supported by Generalitat Valenciana under grant PROMETEO/2019/109 and Spanish Administration and European Union grant TEC2017-84743-P.

II. ALPHA INTEGRATION METHOD

A. Alpha integration for binary classification

Alpha integration was first proposed for the binary classification (detection) problem [7, 8]. Let us assume that we have a group of D binary classifiers (detectors) working on the detection problem. Each detector will produce a score s_i , $i = 1 \dots D$, where higher values of s_i indicate that the positive class is more likely than the negative class. In this context, alpha integration performs the optimal integration of these scores $\mathbf{s} = [s_1 \dots s_D]^T$ into a single score s_α such that

$$s_\alpha(\mathbf{s}) = \begin{cases} \left[\sum_{i=1}^D w_i (s_i)^{(1-\alpha)/2} \right]^{2/(1-\alpha)}, & \alpha \neq 1 \\ \exp\left[\sum_{i=1}^D w_i \log(s_i) \right], & \alpha = 1 \end{cases} \quad (1)$$

where α and the coefficients $\mathbf{w} = [w_1 \dots w_D]^T$ are the parameters to be optimized, subject to $w_i \geq 0$, $\sum_{i=1}^D w_i = 1$. Due to these constraints, s_α is bound between 0 and 1. It can be shown that many classical late soft fusion techniques are particular cases of alpha integration, such as the average ($\alpha = -1$ and $w_i = 1/D \forall D$), the minimum ($\alpha = \infty$) and the maximum ($\alpha = -\infty$). In practice, there are many applications where the parameters of alpha integration are unknown beforehand and have to be estimated from some training data. Previous works have presented the derivations required to optimize alpha integration with respect to the least mean squares (LMSE) and the minimum probability of error (MPE) criteria [8].

B. Alpha integration for multiclass classification

Alpha integration was recently generalized to multiclass classification in a method called separated score integration (SSI) [9]. Essentially, SSI performs alpha integration separately on the scores corresponding to each class.

Given K classes, indexed by $k = 1 \dots K$, and D classifiers, the i th classifier will produce a vector of scores $\mathbf{s}_i = [s_{1i} \dots s_{Ki}]^T$, $i = 1 \dots D$. We will assume the scores are normalized to unit sum, $\sum_{k=1}^K s_{ki} = 1$. The true class identifier vector is defined as $\mathbf{y} = [y_1 \dots y_K]^T$, where

$$y_k = \begin{cases} 1 & \text{if the true class is } k \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Let us define α_k and w_{ki} , $i = 1 \dots D$, as the parameters to integrate the scores corresponding to class k . Given a set of scores $\mathbf{S} = [\mathbf{s}_1 \dots \mathbf{s}_D]$, we can directly apply the integration function (1) to every class

$$s_{\alpha_k}(\mathbf{r}_k) = \begin{cases} \left(\sum_{i=1}^D w_{ki} \cdot s_{ki}^{\frac{1-\alpha_k}{2}} \right)^{\frac{2}{1-\alpha_k}}, & \alpha_k \neq 1 \\ \exp\left(\sum_{i=1}^D w_{ki} \cdot \log(s_{ki}) \right), & \alpha_k = 1 \end{cases} \quad (3)$$

Where $k = 1 \dots K$ and \mathbf{r}_k^T is the k th row of matrix \mathbf{S} . This way, the multi class problem with K classes is converted in K separate two-class problems. The scores are then normalized so that they add up to one:

$$s_{\alpha_k}^{norm} = \frac{s_{\alpha_k}}{\sum_{k=1}^K s_{\alpha_k}} \quad (4)$$

Once we have fused the scores for all classes, classification is performed by selecting the class with the highest score, $\hat{k} = \operatorname{argmax}_k s_{\alpha_k}^{norm} = \operatorname{argmax}_k s_{\alpha_k}$. SSI is an extension of alpha integration of experts in [10] but shares its optimality under alpha risk.

As with alpha integration, the parameters of SSI usually have to be estimated from training data. Derivations to optimize the parameters of SSI with respect to the LMSE and MPE criteria were presented in [9].

III. OTHER FUSION METHODS

In this work, the performance of alpha integration has been compared with two classical fusion techniques: score averaging (mean) and majority voting. Furthermore, it has also been compared with other state-of-the-art methods for late fusion.

A. Copulas

Copulas are advanced methods for the estimation of multivariate probability density functions [11], since Sklar's theorem states that any multivariate pdf can be expressed as the product of a copula function and the product of univariate marginal distributions for each variable. Copulas have been applied successfully in many applications, such as financial model dependence [15].

We used copulas to build a generative classifier using the scores of the single classifiers. Briefly, copulas are used to estimate the multivariate pdf of the scores of all the classifiers for each true class. Then, for every input sample, the posterior probability of each class for that sample is computed using the Bayes theorem, and the sample is assigned to its most likely class (*maximum a posteriori*).

In the following, we considered the t family of copulas and the marginals were estimated using non-parametric kernel density estimation.

B. Independent component analysis mixture model

An alternative to copulas is independent component analysis (ICA), which transforms the input data into a linear combination of independent components. Thus, since the components are independent, their multivariate pdf can also be expressed as a product of univariate marginal distributions. We considered ICA mixture models (ICAMM), which maintain the modeling capabilities of ICA but display increased flexibility [12]. ICAMM has been used in many applications [16-19], including biosignal processing [20]. The estimation of the ICAMM parameters was performed using the MIXCA method [12].

C. Dempster-Shafer combination

Evidence theory or Dempster-Shafer (DS) theory is a general framework for dealing with uncertainty and belief [13]. Dempster-Shafer theory has been used in applications such as fault diagnosis [21].

In this work, we interpreted the scores produced by each classifier as probability masses, and then applied Dempster's rule of combination:

$$(m_1 \oplus m_2)(A) = \frac{1}{1-M} \sum_{B \cap C = A \neq \emptyset} m_1(B) m_2(C) \quad (5)$$

Where $M = \sum_{B \cap C = \emptyset} m_1(B) m_2(C)$, and A , B and C are subsets of the whole universe, with \emptyset being the empty set.

D. Behavior knowledge space

A behavior knowledge space (BKS) is a D dimensional space where each dimension corresponds to the decision of one classifier $\hat{\mathbf{k}} = [\hat{k}_1 \dots \hat{k}_D]^T$ [14]. BKS fusion is a decision fusion method that estimates the posterior probabilities by computing the frequency of each class for every possible set of classifier decisions $\hat{\mathbf{k}}$, based on a given training set. BKS is also understood as a knowledge discovery method [22,23]. BKS has been used in many applications, e.g., detecting copy-move forgery in images [24]. Unlike the previously-presented methods, BKS fuses hard decisions (classes) rather than soft decisions (scores).

IV. EXPERIMENTS

The considered fusion methods were tested on a set of multichannel electroencephalographic (EEG) data from six epileptic subjects that were performing several neuropsychological tests. Evaluating the learning and memory functions of the patients is a critical part of their neuropsychological condition assessment. Each set of EEG signals was captured on 18 bipolar EEG channels set according to the 10-20 system, sampled at 500 Hz. An example of the captured signals is shown in Fig. 1. For pre-processing, a band pass filter (0.5 Hz and 50 Hz) and a notch filter (50 Hz) were applied on each EEG channel. These data have been used in previous works [25].

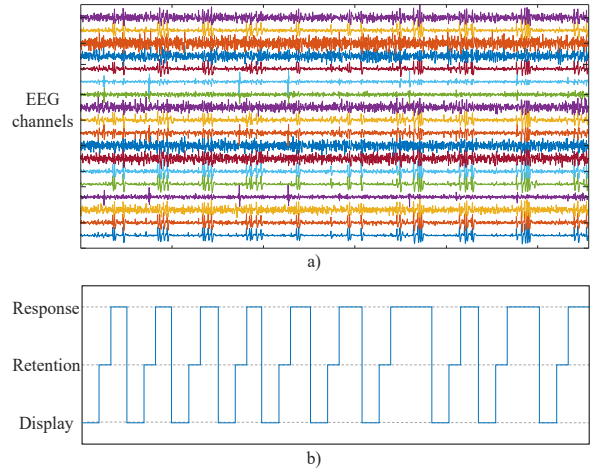


Fig. 1. Example of the captured data for one of the subjects during testing: a) 18 bipolar EEG channels; b) stages of the test.

Each trial of the tests was divided in three stages (classes): stimulus display (SD), retention interval (RI), and subject response (SR). We considered four single classifiers: linear discriminant analysis (LDA); naïve Bayes (NB) with non-parametric kernel density estimation of the marginals; random forests (RDF) using 50 trees; and support vector machines with a linear kernel (SVM). These classifiers were chosen for their successful application in many different applications. The single classifiers were considered in isolation and then

combined using each of the fusion techniques described in Sections II and III: two classical fusion techniques (mean, majority voting); Dempster-Shafer combination; SSI optimized with respect to the LMSE (SSI-LMSE) and MPE (SSI-MPE) criteria; copulas; ICAMM; and BKS.

We considered the following features for classification: average amplitude; average power; centroid frequency; power in frequency bands delta (0-4 Hz), theta (5-7 Hz), alpha (8-12 Hz), sigma (13-15 Hz) and beta (16-30 Hz); and the activity, mobility and complexity of the signal [26]. These features are common in EEG signal processing [27,28] and other classification works [29]. The features were extracted in 1-second epochs with no overlap between epochs. The same features were extracted for each of the 18 EEG channels, resulting in 198 features per epoch. Given the relatively high dimensionality of the data, feature selection was performed before classification. For this task, a feature ranking method was chosen [30]. In this work, the score of each feature was estimated as its average accuracy using a simple classifier that only considered that feature. This accuracy was estimated using 10-fold cross-validation. Then, the features were ranked in descending order of performance and the 10 best-ranked features were chosen.

The number of epochs available for each subject was limited, and the prior probabilities were low for some classes. To test the performance of the system, we performed a set of Monte Carlo experiments. For each iteration of the experiments, we used 10-fold cross-validation. First, we used 10-fold cross-validation to obtain scores of each classifier for every epoch. Then, we used a different 10-fold cross-validation experiment to obtain the scores using the fusion methods that require parameter fitting (alpha integration, BKS, copulas, and ICAMM). This way, every given method was trained using 90% of the samples and then tested on the remaining 10%, keeping as many values as possible for training. Finally, the result was obtained as the average of 100 iterations.

For performance, the kappa coefficient was chosen because of the differences in prior probability for the stages of the test, which might impact performance. The following extension of kappa coefficient for multiclass classification was used [31]:

$$\kappa = \frac{N \cdot \sum_{i=1}^K c_{ii} - \sum_{i=1}^K c_{i+} \cdot c_{+i}}{N^2 - \sum_{i=1}^K c_{i+} \cdot c_{+i}} \quad (6)$$

Where N is the total number of samples; c_{ii} is the number of correctly classified samples that belong to class i ; c_{i+} is the number of samples that belong to class i ; and c_{+i} is the number of samples classified as class i . Unlike other classification performance indicators, such as the accuracy and balanced accuracy, the kappa coefficient accounts for the different prior probabilities of each class and compensates for the possibility of correct classifications occurring by chance. Thus, the kappa coefficient is more complete and demanding than most other classification performance indicators.

A. Test #1: Barcelona test (visual memory)

The first test was the Barcelona test [32]. During each trial of the Barcelona test, the subject is shown a probe item for 3 seconds and, after a 2-second retention interval, the subject must be able to pick the probe item from a set of four similar

items. There are ten trials in the test, with scoring depending on the number of correct items.

The average performance of each method during the experiment is shown in Fig. 2. The best-performing single classifier was RDF, followed by LDA. Out of the considered late fusion techniques, only DS and alpha integration were able to optimally combine the scores from the different classifiers: classical methods and copulas, ICAMM and BKS returned worse results than the best-performing single classifier.

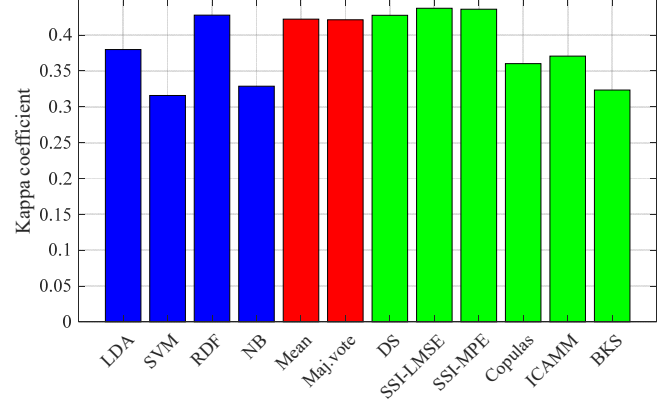


Fig. 2. Average results obtained for the Barcelona test.

B. Test #2: Figural memory

The second test was the figural memory subtest of the Wechsler Memory Scale (WMS-R) [33]. This is an immediate recognition test of abstract designs. The participant is shown three abstract figures for 10 seconds. After a 2-second retention interval, the subject is shown a set of nine similar figures from which they have to select the three figures they were shown before. There are three trials of increasing difficulty, and scoring is calculated from the number of correctly-selected figures.

The average performance of each method during the experiment is shown in Fig. 3. NB and RDF were tied as the best performing single classifiers. With respect to the late fusion methods, the best performance was yielded by ICAMM, which obtained a much better result than the rest of the methods. Alpha integration and BKS were tied as the second best performing late fusion method.

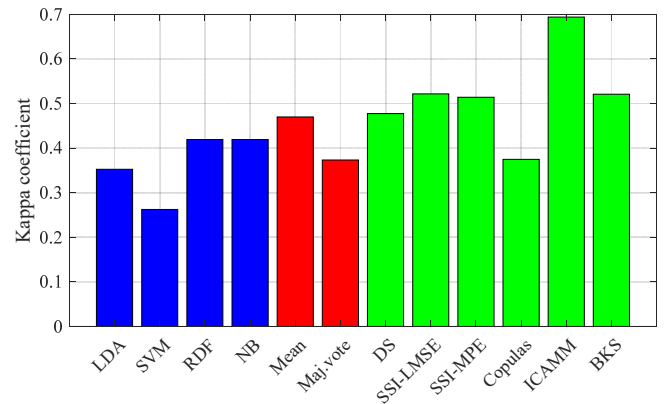


Fig. 3. Average results obtained for the figural memory subtest of WMS-R.

C. Test #3: Visual reproduction

The third test was the Visual Reproduction subtest of the Wechsler Adult Intelligence Scale (WAIS-III) [33]. During

the test, the subject is shown an abstract line figure during 10 seconds. The figure is then removed and, after a 10-second retention interval, the subject must draw the figure from memory. There are three trials of increasing difficulty. Scoring depends on the similarities between the original figure and the reproduction.

The average performance of each method during the experiment is shown in Fig. 4. In this case, RDF was the best performing single classifier, followed closely by NB. As for the figural memory test, the best performing late fusion method was ICAMM. In this case, the second best performance was yielded by copulas.

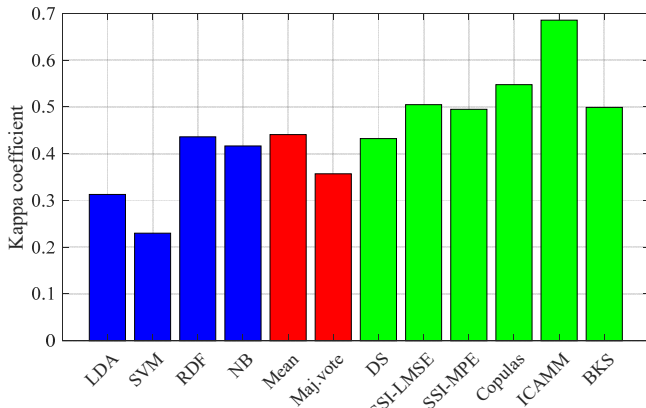


Fig. 4. Average results for the visual reproduction subtest of WMS-R.

D. Test #4: Sternberg memory task

The fourth test was Sternberg’s memory task [34]. During each trial of the task, the subject is shown 2 to 5 probe items (numbers). Each symbol is shown on screen for 0.2 seconds, with a 1-second blank between symbols. Then, after a 1-second retention interval, the subject is shown a test item and asked to determine whether it was one of the probe items. There are 30 trials in the test, with scoring determined by the number of correct responses.

The average performance of each method during the experiment is shown in Fig. 5. The best-performing single classifier was RDF, followed by NB. Out of the considered late fusion techniques, only alpha integration was able to optimally combine the scores from the different classifiers to improve over RDF.

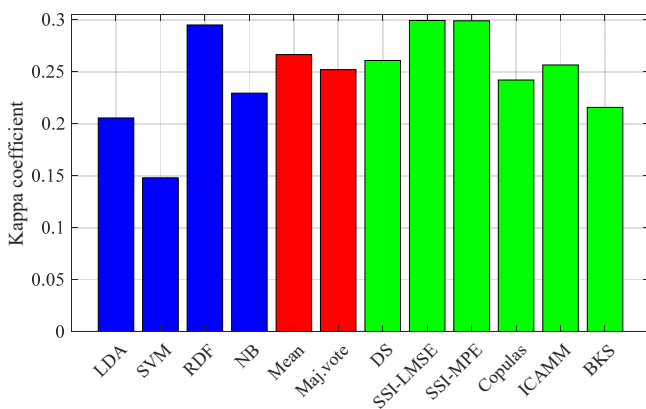


Fig. 5. Average results for Sternberg’s memory task.

V. DISCUSSION

The problem of automatic staging of EEG signals from subjects performing learning and memory tasks is a hard

problem, as shown by the relatively low values of the kappa coefficient of the single classifiers (0.2-0.4, depending on the particular test). In all cases, at least one of the late fusion methods was able to optimally combine the classifiers to obtain improved performance.

The behavior of the late fusion methods was different for each of the considered tests. For instance, BKS and ICAMM oscillated between the best result and a result lower than that of the best single classifier. This performance seemed to be related with the relative performance of the single classifiers. In cases where one of the single classifiers yielded a considerably better result than the rest (see Fig. 2 and Fig. 5), BKS and ICAMM did not yield a good performance. However, when the single classifiers were more comparable (see Fig. 3 and Fig. 4), those late fusion methods yielded optimal performances. Conversely, alpha integration proved to be the more robust late fusion method, always exceeding the performance of the best-performing classifier (RDF or the Gaussian mixture model, depending on the test).

These effects might come from the modeling of the multivariate pdf of the data. In cases where performance was low, the resulting scores were more similar across different classes, and in cases where the number of samples is low (such as this work), the estimated pdfs were not useful. This would also explain the relatively low performance yielded by the copulas. In contrast, alpha integration does not model the pdf of the data, instead fitting a discriminative model for the scores. Thus, it might be more appropriate in cases where we do not have enough samples to properly model the multivariate pdf of the scores.

There are many possible sources of noise in EEG signals, such as line noise and artifacts (eye blinks, muscle activity...) [27]. This noise can make it difficult to find common patterns for the same class, reducing classification performance. Furthermore, the contribution of these sources of noise may not be uniform throughout the tests, with longer and more difficult tests leading to more eye blinks, subject fidgeting that leads to muscle noise, and so on. Noise has not been considered in this work. However, one of the advantages of decision fusion is a reduction of the effect of the noise [3]. Since multiple classifiers are considered, each with its own decision boundary, it is likely that noise will not affect all of them equally and thus their combination will be less affected by said noise. However, more work would be needed in order to confirm this possibility.

VI. CONCLUSIONS

This work has tested the relative performance of several state-of-the-art late fusion methods on a difficult problem with real data: automatically staging electroencephalographic data from subjects performing visual learning and memory tests. The considered late fusion methods were: alpha integration extended to the multiclass case (SSI); copulas; Dempster-Shafer combination (DS); independent component analysis mixture models (ICAMM); and behavior knowledge space (BKS).

The considered fusion methods were employed to combine four classifiers on data from four different neuropsychological tests. These classifiers were trained to perform automatic staging of the tests into three classes: stimulus display, retention interval, and subject response. The considered late fusion methods were able to improve the performance for the automatic staging task. Out of the

considered late fusion methods, ICAMM and BKS yielded very good results in some tests but bad results in others. These behaviors owed to the reduced number of samples available for the problem, which could discourage the estimation of complex multivariate pdfs. Conversely, alpha integration proved to be the most stable, always outperforming the best-performing single classifier.

REFERENCES

- [1] S. Yuksel, J. Wilson, and P. Gader, "Twenty years of mixture of experts," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, pp. 1177-1193, 2012.
- [2] B. Khaleghi, A. Khamis, F. Karray, and S. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion*, vol. 14, pp. 28-44, 2013.
- [3] M. Mohandes, M. Deriche, and S. Aliyu, "Classifiers Combination Techniques: A Comprehensive Review," *IEEE Access*, vol. 6, pp. 19626-19639, 2018.
- [4] S. Wang, V. Anugu, T. Nguyen, N. Rose, J. Burns, M. McKenna, N. Petrick, and R. Summers, "Fusion of Machine Intelligence and Human Intelligence for Colonic Polyp Detection in CT Colonography," in *International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 160-164, Chicago, IL (USA), 2011.
- [5] F. Putze, S. Hesslinger, C.Y. Tse, Y. Huang, C. Herff, C. Guan, and T. Schultz, "Hybrid fNIRS-EEG based classification of auditory and visual perception processes," *Frontiers in Neuroscience*, vol. 8, art. no. 373, 2014.
- [6] J. Zhang, Y. Wu, J. Bai, and F. Chen, "Automatic sleep stage classification based on sparse deep belief net and combination of multiple classifiers," *Transactions of the Institute of Measurement and Control*, vol. 38, no. 4, pp. 435-451, 2015.
- [7] S. Amari, "Integration of stochastic models by minimizing α -divergence," *Neural Computation*, vol. 19, 2796-2780, 2007.
- [8] A. Soriano, L. Vergara, A. Bouziane, and A. Salazar, "Fusion of scores in a detection context based on alpha-integration," *Neural Computation*, vol. 27, 1983-2010, 2015.
- [9] G. Safont, A. Salazar, and L. Vergara, "Multiclass alpha integration of scores from multiple classifiers," *Neural Computation*, vol. 31, no. 4, pp. 806-825, 2019.
- [10] S. Amari, *Information Geometry and its Applications*. Springer, 2016.
- [11] R.B. Nelsen, Roger, *An Introduction to Copulas*. Springer, 1999.
- [12] A. Salazar, *On Statistical Pattern Recognition in Independent Component Analysis Mixture Modelling*. Springer, 2013.
- [13] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [14] Y.S. Huang and C.Y. Suen, "A method of combining multiple experts for the recognition of unconstrained handwritten numerals," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 1, pp. 90-94, 1995.
- [15] D.H. Oh and A.J. Patton, "Modeling dependence in high dimensions with factor copulas," *Journal of Business & Economic Statistics*, vol. 35, no. 1, pp. 139-154, 2017.
- [16] A. Salazar, J. Igual, L. Vergara, and A. Serrano, "Learning hierarchies from ICA mixtures," In *Proceedings of the IEEE International Joint Conference on Artificial Neural Networks*, pp. 2271-2276, 2007.
- [17] G. Safont, A. Salazar, A. Rodriguez, and L. Vergara, "On Recovering missing ground penetrating radar traces by statistical interpolation methods," *Remote Sensing*, vol. 6, no. 8, pp. 7546-7565, 2014.
- [18] A. Salazar, J. Igual, G. Safont, L. Vergara, and A. Vidal, "Image applications of agglomerative clustering using mixtures of non-Gaussian distributions", in *proceedings of the International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 459-463, Las Vegas, NV (USA), 2015.
- [19] A. Salazar, L. Vergara, "ICA mixtures applied to ultrasonic nondestructive classification of archaeological ceramics," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, art. no. 8, 2010.
- [20] G. Safont, A. Salazar, L. Vergara, E. Gomez, and V. Villanueva, "Probabilistic Distance for Mixtures of Independent Component Analyzers," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1161-1173, 2018.
- [21] K.H. Hui, M.H. Lim, M.S. Leong, and S.M. Al-Obaidi, "Dempster-Shafer evidence theory for multi-bearing faults diagnosis," *Engineering Applications of Artificial Intelligence*, vol. 57, pp.160-170, 2017.
- [22] O. Maimon, and L. Rokach, *Data mining and knowledge discovery handbook*. Springer, 2005.
- [23] A. Salazar, J. Gosalbez, I. Bosch, R. Miralles, and L. Vergara "A case study of knowledge discovery on academic achievement, student desertion and student retention," in *proceedings of the 2nd International Conference on Information Technology: Research and Education (ITRE)*, pp. 150-154, London (UK), 2004.
- [24] A. Ferreira, S.C. Felipussi, C. Alfaro, P. Fonseca, J.E. Vargas-Muñoz, J.A., dos Santos, and A. Rocha, "Behavior knowledge space-based fusion for copy-move forgery detection," *IEEE Transactions on Image Processing*, vol. 25, no. 10, pp. 4729-4742, 2016.
- [25] G. Safont, A. Salazar, L. Vergara, E. Gómez, and V. Villanueva, "Multichannel dynamic modeling of non-Gaussian mixtures," *Pattern Recognition*, vol. 93, pp. 312-323, 2019.
- [26] J. Hjorth, "The physical significance of time domain descriptors in EEG analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 34, no. 3, 321-325, 1973.
- [27] S. Sanei, J.A. Chambers, *EEG Signal Processing*. John Wiley & Sons, 2013.
- [28] R. Llinares, J. Igual, A. Salazar, and A. Camacho, "Semi-blind source extraction of atrial activity by combining statistical and spectral features," *Digital Signal Processing: A Review Journal*, vol. 21, no. 2, pp. 391-403, 2011.
- [29] J. Igual, A. Salazar, G. Safont, and L. Vergara, "Semi-supervised Bayesian classification of materials with impact-echo signals," *Sensors*, vol. 15, no. 5, pp. 11528-11550, 2015.
- [30] U. Stańczyk, L.C. Jain, *Feature Selection for Data and Pattern Recognition*. Springer, Berlin (Germany), 2011.
- [31] J. Carletta, "Assessing agreement on classification tasks: The kappa statistic," *Computational Linguistics*, vol. 22(2), pp. 249-254, 1996.
- [32] M. Quintana, J. Pena-Casanova, G. Sánchez-Benavides, K. Langohr, R. Manero, M. Aguilar, D. Badenes, J. Molinuevo, A. Robles, M. Barquero, and C. Antúnez, "Spanish multicenter normative studies (Neuronorma project): norms for the abbreviated Barcelona Test," *Archives of Clinical Neuropsychology*, vol. 26(2), pp. 144-157, 2010.
- [33] E. Strauss, *A Compendium of Neuropsychological Tests*. Oxford University Press, 2006.
- [34] S. Sternberg, "High-speed scanning in human memory," *Science*, vol. 153, no. 3736, pp. 652-654, 1966.