Fuzzy classifiers for chemical compound recognition from SAW sensors signals

Edwin Friedmann, Jean-Philippe Poli, Olivier Hotel, Christine Mer-Calfati

Fuzzy Classifiers for Chemical Compound Recognition from SAW Sensors Signals

Edwin Friedmann
CEA, LIST
91191, Gif-sur-Yvette cedex, France
edwin.friedmann@cea.fr

Jean-Philippe Poli
CEA, LIST
91191, Gif-sur-Yvette cedex, France
jean-philippe.poli@cea.fr

Olivier Hotel
CEA, LIST
91191, Gif-sur-Yvette cedex, France
olivier.hotel@cea.fr

Christine Mer-Calfati
CEA, LIST
91191, Gif-sur-Yvette cedex, France
christine.mer@cea.fr

***Abstract*—Chemical vapor analysis devices are booming, thanks to a growing need in areas such as security and quality control. These control devices are based on various technologies that are the subject of important researches in an ever-growing community of physicists and electronics. However, the data from these sensors are often processed by conventional algorithms poorly configured for the purpose of automatically recognizing target chemical compounds. These algorithms are often based on statistical models that are not always adapted to a limited number of learning data and demonstrated reproducibility problems for these kind of sensors. In this article, we propose to train fuzzy models and compare their performances with the classical methods of the state of the art, to show how practical they can be for such applications. Three different uses cases will be studied: toxic chemicals recognition, detection of counterfeit coffee capsules and detection of a chemical weapon among everyday products.**

***Index Terms*—Fuzzy rule-based system, volatile chemical compounds, surface acoustic wave sensors, FURIA, fuzzy decision trees.**

I. Introduction

For decades, the recognition of volatile chemical compounds was done using classic olfactometry approaches: for instance, using a panel of experts or exploiting the developed olfactory capabilities of certain animals. However, physiological differences and concentration problems led to imprecise, even non-reproducible identifications. For example, it is estimated that a trained dog can only stay focused between 15 and 30 minutes. To overcome these shortcomings, alternative methods have been developed: gas chromatography, mass spectrometry, optical spectrometry and chemical sensors [1].

The problem of classification of "odours" has been successfully treated, in particular with fuzzy logic [2]–[7]. To carry out this task, different descriptors are used for classification, such as the physico-chemical quantities of the molecules [3], [5], the characteristics resulting from the responses of sensory sensors [4], [7], or even sensory analyses carried out by experts [2]. Although they all rely on the formalism of fuzzy logic, the classification methods used are also very varied, ranging from neural networks like Fuzzy ARTMAP [4], to methods of fuzzy rules induction [2], [5] and modeling by experts [3], [7].

This article focuses on chemical sensors, which are booming thanks, among other things, to their transportability and their low cost. Their general principle is based on the physico-chemical interaction between a volatile compound and a layer of sensitive material. In particular, this work deals with SAW (Surface Acoustic Wave) sensors that will be introduced in section II.

In [8], the recognition of chemical compounds has been studied with classical approaches of machine learning (neural networks, SVP, kNN, etc.). However, the reactions of SAW sensors are not exactly reproducible, which leads to the need of a large number of training examples.

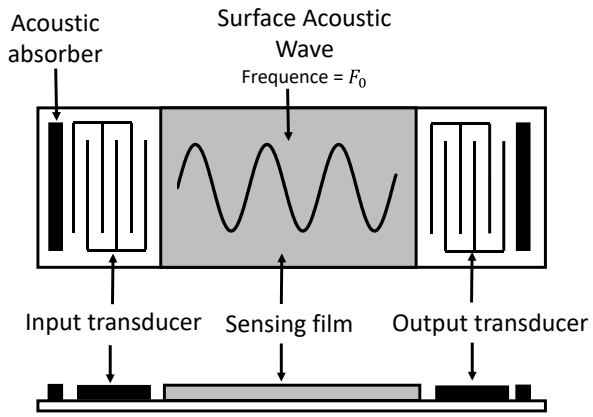
This article addresses the same problem but this time using fuzzy rule-based classifiers. They have the advantage to return a rule-based model that can be evaluated by experts of the domain afterwards. Furthermore, one can have access to the details of rules activated during a process of classification, making the decision more transparent.

As experts cannot formulate a rule base, we use machine learning. We will therefore compare two approaches: FURIA and fuzzy decision trees. These two algorithms represent two different strategies to induce rules from data.

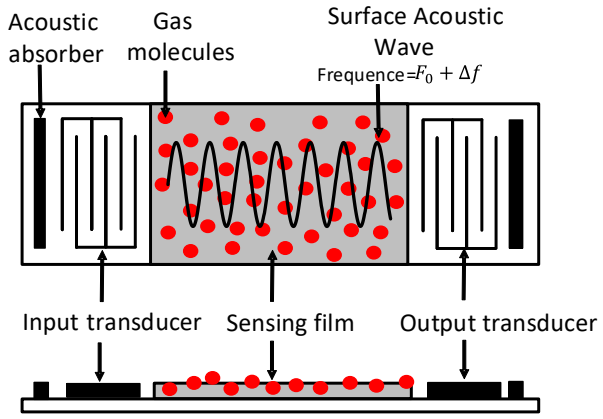
The article is composed as follows. Section II explains the principles of SAW sensors and the feature extraction. Section III reminds the principles of the two methods of rule induction, which are used in this work. We then describe the experimental protocol in section IV and we present the results (section V) and discuss them (section VI) before concluding.

II. SAW Sensors Feature Extraction

The particularity of this work is to be applied on a more recent generation of chemical sensors based on Surface Acoustic Wave (SAW). Figure 1 presents an illustration of such sensors.



(a) The SAW sensor before any volatile chemical compound



(b) The SAW sensor during an exposition to volatile chemical compounds

Fig. 1: Principles of SAW sensors

When molecules interact with the sensitive surface of a SAW sensor, the physical properties of the latter change. These changes are converted into a signal. In our case, modified diamond nanoparticles are used as sensitive coatings that can be mixed with other nanoparticles to improve their sensitivity to chemical compounds. It is thus common to use a matrix of such sensors with different coatings, and so to manipulate a multi-valued signal.

This differs from most of the previous works cited in the introduction since we do not have a characterization of the molecules but only signals from which we will extract features.

Figure 2 represents the response of the sensor when exposed to a coffee capsule. Each curve is in fact the output of a SAW sensor whose functionalization (the reactive layer) is specified in legend. This figure shows that the multi-valued signal is composed of two parts: a phase called transient regime and a phase called stationary regime. In practice, there is undoubtedly information in the transient mode but the sampling rate of these sensors being rather low (of the

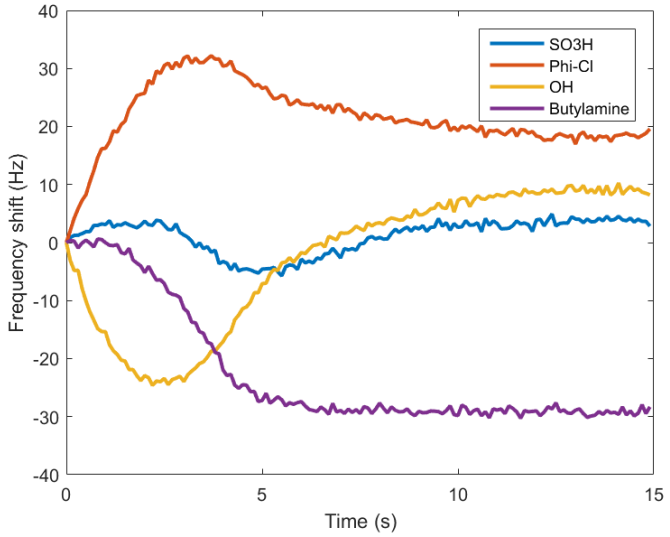


Fig. 2: Example of output of the sensor for a coffee capsule exposure.

order of 10Hz), it often happens that this phase is not seen. The descriptors thus use the stationary phase.

The most used features in the literature [8], [9] are the amplitudes of the different signals during this stationary phase.

However, Hotel et al [10] proposed new interpretable descriptors based on the mathematical equations, which model the vibrations of the reactive layer. Indeed, it was established that the frequency shift of SAW sensors is the superposition of two main contributions (the electro-acoustic one can be neglected) [11]:

1) a viscoelastic contribution due to changes in the coating's Young modulus; and
2) a mass loading effect due to changes of the coating film's mass.

[11] modeled these equations by first order linear differential equations. Since these contributions constitute a signature of the odour, [10] suggests solving the equations with the information contained in the signals. The features are their solutions. Unfortunately, the lack of available information prevent from solving them analytically. [10] uses metaheuristics to approximate the solutions.

These features show a real improvement of the performances in most applications [10]. However, the features represent uncertain data. This motivates the use of fuzzy logic, in particular the two methods we present in the next section.

III. Fuzzy Rule-based Model Induction

A. Learning Fuzzy Rules with FURIA

FURIA [12], standing for *Fuzzy Unordered Rule Induction Algorithm*, is an extension of the RIPPER [13] non-fuzzy rule learning algorithm. Briefly, RIPPER is a multi-label classification algorithm that learns rules having as a premise a conjunction of inequalities on numeric attributes

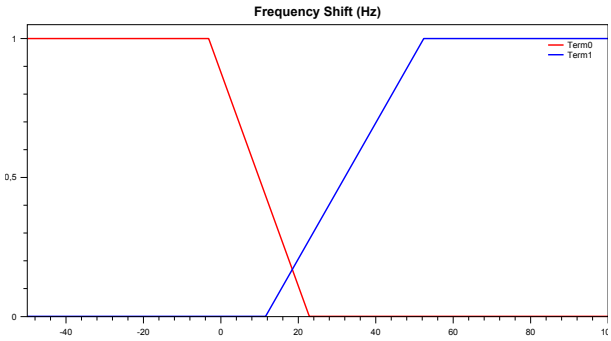


Fig. 3: Example of trapezoidal membership functions used in FURIA

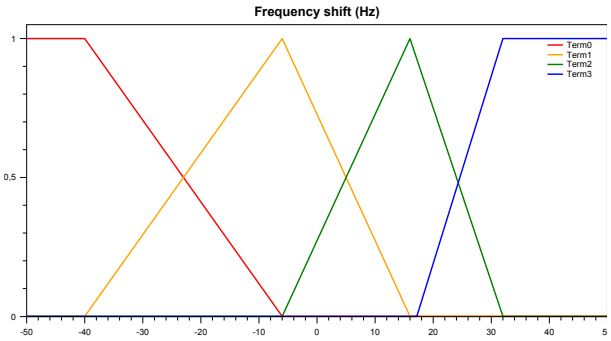


Fig. 4: Strong partition of the domain of an input variable, automatically obtained with a clustering approach

and consequently, a class among the set of classification labels. With FURIA, Hühn and Hüllermeier introduce two major modifications: the learning of unordered rules and the fuzzification of learned inequalities.

The first consists in applying a *one-vs-rest* decomposition of the data to obtain an unordered rule base, instead of eliminating the classes one after the other and get a decision list like previous rule base learning algorithms. The latter approach can cause a bias in favor of the default class, which is corrected in FURIA by treating each class equivalently.

Then, as in RIPPER, the rules for each class are learned by starting with an empty rule and adding the predicates, which maximize the gain of information from FOIL [12] until the rule only covers target data. The rule set stops growing when all the target instances are covered or if the total *description length* of the rule base is 64 bits higher than the smallest previously calculated.

After growth, the previous set of rules is optimized to avoid over-fitting. Each rule R_i is compared to two newly constructed rules: a replacement rule R'_i, empty before growth, and a revision rule R''_i, equal to R_i before growth. They are both temporarily placed in the i^{th} position of the rule set, developed as above according to FOIL's information gain, and pruned to minimize the error of the total rule base.

Between R_i, R'_i and R''_i, the rule that generates the smallest *minimum description length* (MDL) is kept. The remaining target examples are covered as in the growth phase and this optimization can be repeated several times (2 times in general).

The second major modification occurs at the end of the algorithm: the inequalities initially learned are transformed into fuzzy sets with a trapezoidal membership function, as illustrated on Figure 3. For a rule, we determine the best fuzzification of each predicate: the original interval defined at the core of the trapezoid and the data not covered by the predicate are evaluated as the support boundary of the trapezoid, via a purity measurement [12]. The purest of the fuzzy intervals thus constructed is kept and this procedure is repeated until all the inequalities are fuzzified.

Finally, we obtain a set of unordered fuzzy rules. This type of model involves two possible classification problems: examples covered identically by two rules with different conclusions and examples not covered by a rule. For the first, FURIA weighs the rules according to their certainty factor on the learning set; the second is treated by *rule stretching*, i.e. by evaluating by Laplace precision the *minimum generalization* of the rules.

In summary, the FURIA algorithm follows four main steps:

1) One-vs-rest learning of crisp rules by maximizing the gain of FOIL information on predicates.
2) Optimization of each rule by growing 2 alternative rules and choice of the rule generating the smallest *MDL*.
3) Fuzzification of crisp expressions into fuzzy sets with trapezoidal membership functions by maximizing their purity criterion.
4) Classification: weighting of rules and rule-stretching if examples are not covered.

B. Inducing Fuzzy Decision Trees

Fuzzy decision trees (FDT) are a generalization of decision trees [14]. They appear in the early 90s with fuzzified versions of ID3 [15], [16]. To the authors of [17], fuzzy decision trees are classifiers closer to human thinking and that are more robust to imprecision, conflict and missing information.

The most used FDTs are derived from the ID3 algorithm [18]–[20]. They consist in determining strong fuzzy partitioning of the input variables before the learning algorithm in tree structure. The fuzzy sets obtained via clustering (Figure 4) are then used as categorical variables in the original ID3 algorithm, each node dividing a variable into its different fuzzy sets. An instance x is distributed among the children of the node according to its membership degree $\mu_{A_i^j}(x)$ to each fuzzy subset A_i^j of the variable i split by the node. The measure of discrimination, i.e. the entropy for ID3, has been adapted to consider the different membership degrees for each branch.

Another category of fuzzy decision trees groups the algorithms that fuzzify the thresholds used in the nodes of trees handling numerical data [21], [22]. Instead of performing a crisp binary division of a variable, the boundary is fuzzified to obtain a fuzzy transition between the left and right children. In fuzzy SLIQ [22], the width of the linear part of the splitting function depends on the standard deviation of the variable,

which is split. In [21], the authors perform a Fibonacci search to optimize the width of the slope after that the optimal threshold has been found.

More recently, Intuitionistic Decision Trees [23], based on Fuzzy ID3, are showing good performances on several datasets.

In this work, we focus on the fuzzy variants of ID3 and C4.5 decision trees, with the following workflow:

1) Clustering for each input variables in order to get modalities.
2) Strong fuzzy partitioning with triangular membership functions whose tops are the centroids of the clusters.
3) FDT growth: the variable minimizing the fuzzy entropy is selected at each step.
4) Classification: for each leaf, the majority class is selected.

IV. Experimental protocol

The number of sensors can be different from one experiment to another. Each sensor is coated with a diamond nanolayer and has a fundamental frequency of 433.9 MHz [24]–[26]. The signals are recorded and the characteristics on these signals are extracted to constitute the learning and validation sets.

The learning algorithms will be evaluated on three databases consisting of the experiments presented below, and on two types of descriptors presented in the following subsection. The first data set consists of acquisitions of 5 different toxic gases, the second of experiments on authentic and counterfeit coffee capsules, and the third one of acquisitions of DMMP, simulating Sarin gas, everyday products (e.g., soap) and mixtures of compounds.

A. Use Cases

We select three real-world use cases. They are relevant since the experimental environment is less and less controlled: for instance, the toxic chemicals experiment has been conducted in a lab, while the DMMP experiment has been conducted with a backpack and the ambient air. We will now describe the data sets for the three use cases.

1) Toxic chemicals recognition: The device used is based on a network of eight SAW sensors. The sensors were exposed to ammonia (NH_3), sulfur dioxide (SO_2), hydrogen sulfide (H_2S), methanol (CH_3OH) and toluene (C_7H_8) at a concentration of 10 ppm, 8 ppm, 6 ppm, 4 ppm and 2 ppm. The gases were generated from the dilution of a nitrogen calibrated gas. The temperature of the sensors (22^o) and the flow rate (200 ml / min) above them were kept constant. The sensors were exposed to gas for 15 seconds, then the gas cell containing the sensors was purged for 30 seconds. 17 exposure-purge cycles were performed for each gas at each concentration.

2) Detection of counterfeit coffee capsules: This time, the selected device is based on a network of four SAW sensors. The sensors were exposed to 21 types of authentic commercially available coffee capsules and 7 types of counterfeit capsules. The contents of the capsules were emptied into a sealed beaker; the volatile compounds of each sample were transported through the gas cell containing the sensors using a pump. The temperature of the sensors was kept constant (22 ^{o} C). An exposure of 20 cycles (30 seconds) - concentration (30 seconds) was performed for each sample. The measurement process was repeated for 2 different capsules of the same coffee.

3) Detection of DMMP among interferers: The selected device is based on a network of eight SAW sensors. The sensors were exposed to DMMP and interferents: water, ethanol, shower gel and fertilizers, and DMMP-interfering and interfering-interfering mixtures. After diffusion of the gases in a backpack for 10 min, the vapors are inhaled in through the sensors via a pump. The temperature of the sensors was kept constant (22 ^{o} C). Several acquisitions were made, over several days in order to obtain a sufficiently large database.

V. Results

We compare here the performances in terms of classification, obtained by applying statistical learning methods and fuzzy classifiers presented in section III on the three databases and the two types of features: tables I, II and III.

For each data set and each feature set, the performances of the best statistical method are presented [8] (called "baseline" in the tables). As a reminder, they were obtained by a grid search of the best hyperparameters of the SVM, LMNN, BaggedTree and XGBoost methods. In the case of toxic chemicals recognition, the correct classification rate was used as a performance measure, while for the detection of counterfeit coffee capsules and the detection of DMMP among interferers, sensitivity has been calculated.

Similarly, the search for optimal parameters for FURIA and fuzzy decision trees is done via a grid search on the number of rule base optimization for the first (from 0 to 4 iterations), and a grid search on the number of terms learned per input variables for the second (from 3 to 7 terms).

The evaluation is carried out with a 5-fold cross-validation process. The database is divided into 5 subsets of similar size. The first serves as a test base first and the other four as a learning base, and so on, from the second to the fifth. Each table collects the mean and the standard deviation of the evaluation metric. In the case of fuzzy classifiers, the number of average rules is also indicated in order to judge the complexity of the models learned. The duration of the model induction is also shown.

Overall, the cross-validation performance of fuzzy classifiers on the three datasets gives at least results close to those obtained with statistical models, or even better in some cases. Numerically, there is often less than 1 % difference between the two types of methods. This is particularly the case for the sensitivities obtained on the DMMP basis (table III).

On the coffee and DMMP bases, we also find an improvement in the results, tables II and III, when we consider the mass and viscoelastic contributions compared to the results

TABLE I: Accuracy on toxic chemicals

Eight amplitudes

	Score	Rules	Duration
Baseline	$98.1 \pm 0.7\%$	–	$< 1s$
FURIA	$96.4 \pm 1.7\%$	21	$1.2s$
FDT	$98.8 \pm 0.9\%$	203.6	$9.7s$

Sixteen contributions

	Score	Rules	Duration
Baseline	$98.5 \pm 0.8\%$	–	$< 1s$
FURIA	$95.0 \pm 1.8\%$	23.4	$1.0s$
FDT	$98.0 \pm 0.9\%$	219	$15s$

TABLE II: Sensitivity on coffee data set

Four amplitudes

	Score	Rules	Duration
Baseline	$62.6 \pm 8.7\%$	–	$< 1s$
FURIA	$52.6 \pm 1.9\%$	32.4	$2.9s$
FDT	$61.2 \pm 4.0\%$	197.2	$35s$

Eight contributions

	Score	Rules	Duration
Baseline	$69.8 \pm 7.9\%$	–	$< 1s$
FURIA	$69.5 \pm 4.3\%$	18.8	$5.6s$
FDT	$74.2 \pm 3.3\%$	938.6	$110s$

TABLE III: Sensitivity on DMMP data set

Eight amplitudes

	Score	Rules	Duration
Baseline	$96.6 \pm 2.6\%$	–	$< 1s$
FURIA	$96.8 \pm 4.4\%$	8	$0.2s$
FDT	$96.7 \pm 5.0\%$	57.6	$4.6s$

Sixteen contributions

	Score	Rules	Duration
Baseline	$98.3 \pm 1.1\%$	–	$< 1s$
FURIA	$97.7 \pm 3.1\%$	7.6	$0.2s$
FDT	$97.7 \pm 3.2\%$	58.2	$2.0s$

obtained with the amplitudes as descriptors. The gain is most significant on the basis of coffee capsules: +13 % for the FDT algorithm and +16.9 % for FURIA.

VI. Discussion

A. FURIA

The FURIA algorithm gives scores close to statistical learning methods for DMMP, table III, but lower in the case of toxic chemicals, table I. For the recognition of counterfeit coffees, table II, the score is 10% lower with the amplitudes as descriptors but goes back to the level of the statistical methods with the contributions. Apart from the number of rule base optimizations, the scores were obtained with standard values for hyperparameters: the minimum number of instances covered by a premise is fixed to 2 and the number of folds for optimization to 3. If these hyperparameters are varied, there is no significant change in the performances.

In addition to classification performance, we note that the rule bases learned by FURIA are small, from 7 to 32 rules in average. Regarding the number of training examples for each base (600, 1600 and 160 respectively), there is at least one rule for 20 examples in average. Thus, there is no over-fitting with this method and the models obtained are compact.

From the confusion matrices given in table IV, we verify that there is no bias in favor of the most frequent class CH_3OH like it could appear with RIPPER algorithm.

TABLE IV: Confusion matrices on toxic chemicals
Eight amplitudes

Predicted

Actual	CH_3OH	NH_3	SO_2	C_7H_8	H_2
CH_3OH	170	0	0	0	0
NH_3	0	133	1	4	0
SO_2	0	0	158	0	0
C_7H_8	0	1	8	141	0
H_2	0	0	0	0	140

(a) Confusion matrix for baseline

Predicted

Actual	CH_3OH	NH_3	SO_2	C_7H_8	H_2
CH_3OH	168	0	0	2	0
NH_3	0	129	3	5	1
SO_2	0	0	155	1	2
C_7H_8	1	7	1	140	1
H_2	0	2	1	0	137

(b) Confusion matrix for FURIA

Predicted

Actual	CH_3OH	NH_3	SO_2	C_7H_8	H_2
CH_3OH	170	0	0	0	0
NH_3	0	134	0	4	0
SO_2	0	0	158	0	0
C_7H_8	1	3	0	146	0
H_2	0	0	1	0	139

(c) Confusion matrix for FDT

B. Fuzzy Decision Trees

The fuzzy decision trees give in the three cases quasi-equivalent scores to the statistical methods: the difference is at most 1.4 %, except with the contributions on the coffee data set where there is a gain of 4.4 %, table II.

These performances were obtained by testing a few combinations on the hyperparameters of the induction algorithm, namely, the minimum degree of membership of an example in a branch, the termination criteria on the gain of entropy, the sum of the degrees of belonging to a node, and the minimum number of examples in a node. Compared to FURIA, performance varies more strongly by changing the hyperparameters of the induction algorithm.

The strong partitioning of the input variables facilitates the interpretability of the model. With, for example, a partition with three fuzzy sets, we can assign the terms *Low*, *Medium* and *High* to the learned sets, making the rule base accessible to the final users. There is however a greater number of rules learned than for FURIA. We note that for the coffee data set, 938 rules for 1600 examples are induced in average, i.e. less than two examples per rule. The model suffers from over-fitting and the accuracy (or sensitivity) as the only performance criterion is not sufficient. It is then a question of finding a compromise between performance at evaluation, number of rules and computation time. We can also use rule base simplification algorithms to return reduced size models.

VII. Conclusion and perspectives

In this article, we have successfully applied fuzzy rule-based systems to the recognition of chemical compounds by an array of SAW sensors. In order to get the rule bases, we chose to use two popular methods: FURIA and fuzzy decision trees.

We collected data in three different use cases: the detection of toxic compounds, the recognition of counterfeit coffee capsules and the detection of the presence of DMMP in a backpack. We also retained two types of descriptors in the bibliography, based on one hand on the amplitudes, and on the other hand on the mass and the visco-elastic contributions.

With a few exceptions, fuzzy classifiers lag behind or even exceed statistical classifiers, with an accuracy of almost 99% on toxic chemicals. In addition, we can note that the standard deviations remain relatively low with the fuzzy methods, less than 5%, which shows a good robustness of the models in these applications. From all of these results, we conclude in the interest to add these fuzzy classifiers in our scope of search for the model learning process.

The use case of coffee capsules is delicate and the results are no doubt due to the limitations of the sensors used: in fact, coffee is made up of around a thousand different molecules and the functionalization of the sensors were the same for all three use cases.

We intend to continue this work by making better use of the quality of the descriptors based on the contributions. Indeed, coming from an approximation based on a meta-heuristic optimization method, it is possible to fall into local extrema and have degraded solutions. We will study the quality of these features and use it during learning and recognition.

References

[1] S. Li, "Overview of odor detection instrumentation and the potential for human odor detection in air matrices," MITRE Nanosystems Group, Tech. Rep., 2009.

[2] S. Guillaume and B. Charnomordic, "Fuzzy inference systems to model sensory evaluation," in *Intelligent Sensory Evaluation: Methodologies and Applications*, D. Ruan and X. Zeng, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 197–216.

[3] M. Kissi, M. Ramdani, B. Bouchon-Meunier, and D. Zakarya, "Pattern recognition system based on empirical knowledge: Sandalwood and camphoraceous odours application," *Mathematics and Computers in Simulation*, vol. 77, no. 5, pp. 453 – 463, 2008.

[4] E. Llobet, E. L. Hines, J. W. Gardner, P. N. Bartlett, and T. T. Mottram, "Fuzzy artmap based electronic nose data analysis," *Sensors and Actuators B: Chemical*, vol. 61, no. 1, pp. 183 – 190, 1999.

[5] C. Marsala, M. Ramdani, D. Zakarya, and M. Tollabi, "Fuzzy Decision Trees to Extract Features of Odorous Molecules," in *Uncertainty in Intelligent and Information Systems*. World Scientific, 2000, pp. 235–249.

[6] A. Modak, R. B. Roy, B. Tudu, R. Bandyopadhyay, and N. Bhattacharyya, "A novel fuzzy based signal analysis technique in electronic nose and electronic tongue for black tea quality analysis," in *2016 IEEE First International Conference on Control, Measurement and Instrumentation (CMI)*, 2016, pp. 279–283.

[7] B. Szulczyński, J. Gębicki, and J. Namieśnik, "Application of fuzzy logic to determine the odour intensity of model gas mixtures using electronic nose," in *E3S Web of Conferences*, ser. E3S Web of Conferences, vol. 28, Jan. 2018, p. 01036.

[8] O. Hotel, J.-P. Poli, C. Mer-Calfati, E. Scorsone, and S. Saada, "A review of algorithms for saw sensors e-nose based volatile compound identification," *Sensors and Actuators B: Chemical*, vol. 255, pp. 2472 – 2482, 2018.

[9] S. D. P. J. L. W. C. P. J. Yan, X. Guo and S. Zhang, "Electronic nose feature extraction methods: A review," *Sensors*, vol. 15, no. 11, p. 27804–27831, 2015.

[10] O. Hotel, J. Poli, C. Mer-Calfati, E. Scorsone, and S. Saada, "Saw sensor's frequency shift characterization for odor recognition and concentration estimation," *IEEE Sensors Journal*, vol. 17, no. 21, pp. 7011–7018, Nov 2017.

[11] V. B. Raj, A. T. Nimal, Y. Parmar, M. Sharma, and V. Gupta, "Investigations on the origin of mass and elastic loading in the time varying distinct response of zno saw ammonia sensor," *Sensors and Actuators B: Chemical*, vol. 166, pp. 576–585, 2012.

[12] J. Hühn and E. Hüllermeier, "Furia: an algorithm for unordered fuzzy rule induction," *Data Mining and Knowledge Discovery*, vol. 19, no. 3, pp. 293–319, Dec 2009.

[13] W. W. Cohen, "Fast effective rule induction," in *Machine Learning Proceedings 1995*, A. Prieditis and S. Russell, Eds. San Francisco (CA): Morgan Kaufmann, 1995, pp. 115 – 123.

[14] D. Dubois, H. M. Prade, and J. C. Bezdek, *Fuzzy Sets in Approximate Reasoning and Information Systems*. Norwell, MA, USA: Kluwer Academic Publishers, 1999.

[15] K. J. Cios and L. M. Sztandera, "Continuous id3 algorithm with fuzzy entropy measures," in *[1992 Proceedings] IEEE International Conference on Fuzzy Systems*, March 1992, pp. 469–476.

[16] R. Weber, "Fuzzy-ID3: A class of methods for automatic knowledge acquisition," in *2. Int. Conf. on Fuzzy Logic & Neural Networks. Iizuka, Japan 1992*, 1992, pp. 265–268.

[17] Y. Yuan and M. J. Shaw, "Induction of fuzzy decision trees," *Fuzzy Sets and Systems*, vol. 69, no. 2, pp. 125–139, Jan. 1995.

[18] C. Z. Janikow, "Fuzzy decision trees: issues and methods," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 28, no. 1, pp. 1–14, Feb 1998.

[19] W. Pedrycz and Z. A. Sosnowski, "The design of decision trees in the framework of granular data and their application to software quality models," *Fuzzy Sets and Systems*, vol. 123, no. 3, pp. 271 – 290, 2001.

[20] Z. Qin and J. Lawry, "Decision tree learning with fuzzy labels," *Information Sciences*, vol. 172 (1-2), pp. 91 – 129, 6 2005, publisher: Elsevier.

[21] C. Olaru and L. Wehenkel, "A complete fuzzy decision tree technique," *Fuzzy Sets and Systems*, vol. 138, no. 2, pp. 221 – 254, 2003.

[22] B. Chandra and P. P. Varghese, "Fuzzifying gini index based decision trees," *Expert Systems with Applications*, vol. 36, no. 4, pp. 8549 – 8559, 2009.

[23] P. Bujnowski, E. Szmidt, and J. Kacprzyk, "An approach to intuitionistic fuzzy decision trees," in *IFSA-EUSFLAT*, 2015, pp. 1253 – 1260.

[24] E. Chevallier, E. Scorsone, and P. Bergonzo, "New sensitive coating based on modified diamond nanoparticles for chemical saw sensors," *Sensors and Actuators B: Chemical*, vol. vol. 154, no. 2, pp. p. 238–244, 2011.

[25] H. Girard, J. C. Arnault, S. Perruchas, S. Saada, T. Gacoin, J. P. Boilot, and P. Bergonzo, "Hydrogenation of nanodiamonds using mpcvd: A new route toward organic functionalization," *Diamond and Related Materials*, vol. vol. 19, no. 7, pp. p. 1117–1123, 2010.

[26] H. Girard, S. Perruchas, C. Gesset, M. Chaigneau, L. Vieille, J. C. Arnault, P. Bergonzo, J. P. Boilot, and T. Gacoin, "Electrostatic grafting of diamond nanoparticles: a versatile route to nanocrystalline diamond thin films," *ACS applied materials & interfaces*, vol. vol. 1, no. 12, pp. p. 2738–2746, 2009.