



Correction of Belief Function to Improve the Performances of a Fusion System

Didier Coquin^(✉), Reda Boukezzoula^(ID), and Rihab Ben Aneur

LISTIC, University of Savoie Mont Blanc, Annecy le vieux, 74944 Annecy, France
didier.coquin@univ-smb.fr

Abstract. Our application concerns the fusion of classifiers for the recognition of trees from their leaves, in the framework of belief functions theory. In order to improve the rate of good classification it is necessary to correct Bayesian mass functions. This correction will be done from the meta-knowledge which is estimated from the confusion matrix. The corrected mass functions considerably improve the recognition rate based on the decisions provided by the classifiers.

Keywords: Belief functions theory · Mass correction · Meta-knowledge · Fusion of classifiers

1 Introduction

Tree species recognition is the problem of identifying the species of a given tree. This task may be easy for a botanist who has strong knowledge about trees. However, a novice or a lover of the trees universe may have difficulties. This paper is part of the ReVeRiES project that seeks to develop a mobile application that recognizes tree species without needing an internet connection. Indeed, the application will be used in nature, forests, mountains where an internet connection is not available. In the context of tree species recognition, we treat a real problem since the photos of leaves are taken in the natural environment. The processing of this type of data is complicated because of their specificities due firstly to the nature of the objects to be recognized (inter-species similarity and intra-species variability) and secondly to the environment. Errors can be accumulated during the pre-fusion process. The merit of the fusion is to take into account all the imperfections that can taint the available data and try to model them well. The fusion is more effective if the data are well modeled. The theory of belief functions represents one of the best theoretical frameworks able to manage and represent uncertainty, inaccuracy, conflict, etc. This theory is important because of its wealth of tools to manage the various sources of imperfections as well as the specificities of the available data. In the framework of this theory, it is possible to model the data through the construction of mass functions. It is also possible to manage the computational complexity thanks to the approximations allowing to reduce the number of focal elements. Conflict

being one of the most present sources of imperfections, can be dealt through the selection of the best combination rule. By merging sources of information with different degrees of reliability, the least reliable source may affect the data issued from the most reliable one. One of the solutions for this problem is to try to improve the performances of the least reliable source. Thus, by merging with other sources, it will provide useful information and will in turn contribute in improving the performance of the fusion system. The performance improvement of an information source can be affected through the **correction of mass functions**. In this context, the correction can be made based on measures of the relevance or sincerity of the studied source [14]. The confusion matrices present a data source from which meta-knowledge characterizing the state of a source can be extracted [17]. In this paper, the proposed fusion system is a hierarchical fusion system set up within the framework of belief function theory. It allows to merge data from leaves and provides the user with a list of the most likely species while respecting the educational purpose of the application. The computational complexity of this fusion system is quite small allowing, in the long term, to implement the application on a Smartphone.

The paper is organized in the following way: Sect. 2 describes the fusion system. Section 3 presents the different mechanisms for correcting the mass functions. A correction method is proposed in Sect. 4 which we have applied to our fusion system. Section 5 shows the improvement brought by this correction on real data.

2 Fusion System

The sub-classification strategy consists on recognizing a leaf through its botanical properties. The application consists of taking a photo of a leaf with a smartphone and the system extracts morphological characteristics (the lobe shape $\overrightarrow{A_F}$ described by the Polygonal model, the Margin $\overrightarrow{A_C}$, the Base and the Apex $\overrightarrow{A_{BS}}$). The fusion system is composed of three classifiers based on random forests that have been previously trained on learning images. As outputs of each classifier S_i , we have a distribution of probabilities $P_j[e_i]$ in the species referential with $e_i \in \Omega = 72$ the number of species [1]. The fusion system, shown in Fig. 1, consists of 3 steps:

1. **Mass functions construction:** the passage towards the belief functions theory requires the construction of mass functions based on the distributions of probabilities.
2. **Mass functions approximation:** to reduce the computational complexity.
3. **Mass functions combination:** in the framework of the belief functions theory, several combination rules exist. Each one allows to manage one of more sources of imperfections.

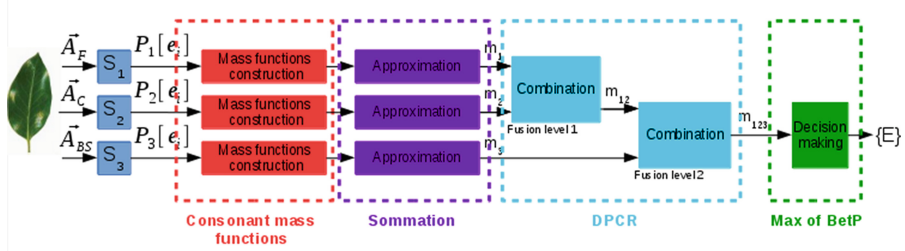


Fig. 1. Hierarchical fusion system for combining leaves data characteristics

2.1 Mass Functions Construction

Let $\Omega = \{w_1, w_2, \dots, w_N\}$ be a frame of discernment composed of N exclusive and exhaustive hypothesis w_i . The set of all subsets A of Ω is called the power set of Ω and is denoted $2^\Omega = \{A | A \subseteq \Omega\} = \{\emptyset, w_1, w_2, \dots, w_N, w_1 \cup w_2, \dots, \Omega\}$. A mass function on Ω is a mapping m from the power set 2^Ω , to the interval $[0, 1]$ such that:

$$\sum_{A \subseteq \Omega} m(A) = 1 \tag{1}$$

Subsets A of Ω such that $m(A) > 0$ are called *focal sets* of m . A mass function is said to be: *vacuous* if Ω is its only focal set, in which case it is denoted by m_Ω ; *inconsistent* if \emptyset is its only focal set, in which case it is denoted by m_\emptyset ; *dogmatic* if Ω is not a focal set; *normal* if \emptyset is not a focal set.

Mass functions construction from probability distributions is an important step in the fusion process. Indeed, the mass functions must be sufficiently representative of all available information. In our case, the data obtained at the output of the classifiers are uncertain because the sources have different degrees of reliability and are sometimes conflicting. It is therefore important that the mass functions can represent all these imperfections and take into account as much information as possible [1]. In our case, we have a large discernment framework ($\Omega = 72$ assumptions). To construct 2^{72} focal elements and to give a mass to all possible combinations would lead to significant computational complexity. In the literature, there are several methods for constructing a mass function from a probability distribution, the best known of which are: bayesian mass functions and consonant mass functions [2]. We have studied and analyzed these two methods in depth. Following these studies, we found that by building bayesian mass functions, we obtain precise and certain focal elements. The obtained mass functions are representative enough of the available information and the focal elements are singletons which causes the elimination of certain species. Concerning consonant mass functions, we noticed that this method leads to the construction of large focal elements (made up of several species). It has the advantage of allowing a better presentation of all available information.

2.2 Mass Functions Approximation

One of the major problems of belief function theory is the computational complexity which is relatively high compared to other theories such as probability theory and possibility theory. Computational complexity increases exponentially with the number of sources to be combined, and becomes even more important when the space of discernment is large. To limit the complexity of calculations, existing approaches rely either on procedures that involve exact calculations or on approximation techniques. In the first category, an optimal calculation algorithm has been proposed by Kennes in [9]. This algorithm makes it possible to considerably optimize the number of addition and multiplication operations when calculating mass functions and to reduce the complexity of the calculations. The disadvantage of this method is that the number of focal elements is not reduced since it is a procedure that performs exact calculations. This can be problematic in our case because we have to manage a large area of discernment and perform the same calculation for several sources of information. The second category of approaches is composed of mass functions approximation procedures. The mass functions approximation procedures allow the reduction of the number of focal elements based on different selection criteria. In particular, we focused on the most commonly used procedures: the “Inner” and “Outer” approximations [4], the $k - l - x$ [21] and the “Summation method” [11]. The analysis, as well as the experiments carried out, have led to the choice of the summation approximation as the most suitable method for our case since it offers the best relationship between computational complexity and preservation of a maximum of useful information.

“**Summation**” approximation consists of keeping the $k - 1$ focal elements having the largest mass. The union of the remaining focal elements forms the k^{th} focal element whose sum of their masses is transferred to the element k . The application of this approximation in our specific context is interesting. Unlike the $k - l - x$ approximation, even focal elements with low mass will be kept. Thus, a maximum of information will be transferred allowing more flexibility to the fusion process. In addition, this approximation is simple, and it does not require a lot of calculation operations which reduces the processing time.

2.3 Mass Functions Combination

To better manage the imperfections of the data, several combination rules have been proposed as part of the theory of belief functions. Each rule has been put in place to handle one or more aspects of imperfections. Given the limited and variable performance of the data sources to be combined, we consider that we have to deal with unreliable sources of data. According to Lefevre et al. [10] the combination rules proposed in the literature represent solutions to manage the different sources of conflict. They divide these combination rules into two categories:

1. the first category groups together the combination rules requiring that the sources of information to combine are reliable.

2. the second category groups combination rules where at least one of the sources of information is reliable without necessarily knowing which one.

In our case, the main problematic situation we face is the following:

- no intersection between 2 focal elements issued from two sources. In that case, our need is to transfer mass to the elements involved in the conflict.
- two or more sources propose the true species. In that case, our need is to transfer mass to the intersection.
- no intersection between the species proposed by the different sources (total conflict). In that case, our need is to transfer mass to the partial ignorance.

After the study and analysis of the different existing combination rules: Dempster combination rule [18], Dubois and Prade combination rule [6], Mixte combination rule, Florea and Jousselme combination rule [8], PCR6 [12] and DPCR combination rule [13], we found that the **DPCR combination rule** is the best compromise between a good recognition rate of the first species and the processing time per species (see Table 1). Indeed, it transfers the mass to the intersection, the elements involved in the conflict as well as to partial ignorance. It also offers the best compromise between classification ratio and computational complexity. DPCR is called the rule of proportional redistribution of the weakened conflict, this rule was proposed by Martin [13]. It is an extension of the *PCR5* combination rule. It consists in applying a discounting procedure to transfer part of the partial conflict to partial ignorance (the union).

Table 1. Processing time per species and recognition rate according to the combination rule

Combination rule	Processing time per species	Recognition rate
Florea	522 ms	57.1%
PCR6	163 ms	51%
Dubois et Prade	199 ms	57.2%
Mixte	634 ms	56%
DPCR	305 ms	58.3%

2.4 Decision Making Criteria

Since the objective of the application is to present a list of the top 10 most probable species, we are interested in the decision criterion “the maximum of pignistic probability” or BetP which is proposed by Smets in [20]. These decision-making criteria makes it possible to transform the mass functions into measures of probabilities for all $w_k \in \Omega$, but in our case, we work with bayesian bba and in a such a case, we have the same results. Thus, we can to provide the user with an ordered list of the most likely species and not an ordered list of the most

likely sets of species. The criterion makes it possible to lift the inaccuracy which can taint the data since we work with consonant mass functions. The dashed colored lines in the Fig. 1 represents the different choices made for the different blocks of the fusion system, which have been previously explained. The three sources S_i are three feature vectors (the Polygonal model, the Margin, the Base and the Apex) that are used as input for the three classifiers (Random Forest).

3 Belief Function Corrections

The theory of belief functions provides a flexible framework for modeling and managing information uncertainty. Several tools are proposed in the literature, they allow to modify or correct some of the information based on additional information on the relevance or sincerity of a source.

3.1 Discounting Based on Degree of Relevance

Historically, discounting is the first correction tool proposed in the literature [19]. It is based on the degree of relevance of an information source S to weaken the mass of information it provides [16]. The discounting was developed as part of the Transferable Belief Model (TBM) [20]. Suppose that an information source S provides information represented by a mass function m_S and assume that β with β in $[0, 1]$, is the degree of relevance relating to this source S . The discounting is done as follows:

$$m(A) = \beta m_S(A) + (1 - \beta) m_\Omega(A), \forall A \subseteq \Omega \quad (2)$$

Discounting therefore consists of weakening the initial mass function according to its degree of relevance and giving more weight to ignorance. The reverse operation of the discounting, called, de-discounting was introduced in [5]. If after making the discounting it turns out that the information on the relevance of the source was false, the de-discounting allows to find $m_S(A)$ from $m(A)$. This correction method uses the overall degree of relevance to weaken the mass functions. However, this measure is insufficient since the degree of relevance may vary from one class to another. Given the quality of the data, it would be more useful to apply contextual correction methods where the weakening of a mass function takes into account each context.

3.2 Contextual Discounting Based on the Degree of Relevance

In [15], Mercier et al. studied contextual discounting based on the degree of contextual relevance of an information source. Its application requires the presence of additional information on the relevance of an information source conditionally in different contexts \mathcal{A} of Ω such as the union of subsets \mathcal{A} form a partition of Ω . Suppose that β_A , with β_A in $[0, 1]$, is the degree of relevance of a source S in a context A and the union of context elements \mathcal{A} form the \mathcal{A} partition of Ω .

The belief m obtained by taking contextual elements \mathcal{A} into account is defined as follows:

$$m = m_S \bigcirc_{A \in \mathcal{A}} A_{\beta_A} \tag{3}$$

where $\bigcirc_{A \in \mathcal{A}}$ represents the canonical disjunctive decomposition, A_{β_A} represents a simple mass function having as focal elements \emptyset and A whose corresponding masses are β_A and $1 - \beta_A$ respectively. A more general case of contextual discounting is discussed in [14]. Its application is simpler and more general since the formation of a partition by A context elements is no longer required. Thus, this correction tool can be applied even if the A context elements do not form a partition of Ω . Suppose β_A , with β_A the degree of relevance given to a source of information in a context A and that \mathcal{A} is the set of contexts for which we know contextually the degree of relevance of the information source. The resulting mass function by taking into account these context elements is calculated as follows:

$$m = m_S \bigcirc_{A \in \mathcal{A}} \{ \bigcap_{A \in \mathcal{A}} (\overline{A})^{1-\beta_A} \} \tag{4}$$

where $\bigcap_{A \in \mathcal{A}}$ represents the canonical conjunctive decomposition. The contextual discounting discussed in this section is based on contextual elements that provide information about the contextual relevance of the information source. This technique of mass function correction is interesting in our case since the relevance of our classifiers is variable from one species to another. With this technique, it would be possible to introduce additional information on the contextual relevance of each source of information to each species. Contextual discounting can also be applied to contextually modify a mass function based on another measure of reliability: sincerity.

3.3 Correction Techniques Based on the Degree of Sincerity

a) Discounting based on the degree of sincerity

It is possible to estimate the sincerity or not of a source of information by assuming that it is reliable. Thus, the correct decisions and incorrect decisions of the source respectively correspond to a sincere or insincere behavior of the source of information. Suppose a source of information is sincere with a β degree and not sincere with a $1 - \beta$ degree. According to [7], the correction of mass functions by taking into account meta-knowledge about the sincerity of a source can be done in 3 different ways:

- by weakening the initial mass while giving mass to ignorance as follows:

$$m = \beta m_S + (1 - \beta) m_\Omega \tag{5}$$

- by weakening the initial mass given to A ($\forall A \subseteq \Omega$) while tuning more mass to its complement \overline{A} ($\forall \overline{A} \subseteq \Omega$) as follows:

$$m = \beta m_S + (1 - \beta) \overline{m_S} \tag{6}$$

- by weakening the initial mass by a factor of $\beta = 1 - 2\epsilon$, where ϵ is the classification error rate in the previous Equation, and guarantying mass to ignorance.

According to the degree of sincerity of the source of information, the initial mass is weakened while granting mass to ignorance as shown by the Eqs. 5. The Eq. 6 presents a slightly different way since the weakening of the initial mass is coupled to a mass transfer to \bar{A} which is the complement set of the focal element A in question. This approach is very interesting because it allows a better preservation of the specificity of the information. In the next section, we are particularly interested in this approach and will propose a method to contextually define the complement set \bar{A} of A .

b) The contextual discounting based on the degree of sincerity

The definition of different forms of non-veracity in [17] made it possible to define a new form of contextual discounting based on the degree of sincerity. Suppose the source S is negatively untrue in \bar{A} . This new form of contextual discounting is defined as follows:

$$m = m_S \bigcirc_{A \in \mathcal{A}} A_{\beta A} \quad (7)$$

c) Contextual reinforcement based on the degree of sincerity

In [14], Mercier and et al. define a new correction mechanism called contextual reinforcement. Suppose the source S is positively untrue in A . This new form of contextual reinforcement is defined as follows:

$$m = m_S \bigoplus_{A \in \mathcal{A}} A_{\beta A} \quad (8)$$

These two new forms of discounting represented in the Eqs. 7 and 8 are interesting because the discounting takes into account the contextual sincerity of the source of information. Its application requires the setting up of assumptions in relation to the sincerity of the source. So we need meta-knowledge to assert if the source is negatively untrue in A or if it is positively untrue in \bar{A} .

3.4 Estimation of Knowledge from a Confusion Matrix

The application of the correction mechanisms requires knowledge of the state of the information source to be processed. This knowledge, called “meta-knowledge”, is used to characterize the state of a data source: relevant or not, sincere or not. This meta-knowledge is generally uncertain and often presented by mass functions [17]. In the context of classifiers fusion, we are interested in the estimation of meta-knowledge from the confusion matrices obtained for each classifier. A confusion matrix $M_i = (n_{kl})_{k \in \{1, \dots, K\}, l \in \{1, \dots, K\}}$ is a table characterizing the performance of a data source S_i on a test set of K classes (Table 2). Each line k in the confusion matrix represents the decision made in favor of w_{lk} . Each column l represents the truth w_l . Note that $n_{k.} = \sum_{l=1}^K n_{kl}$ represents the number of objects in w_k and $n_{.l} = \sum_{k=1}^K n_{kl}$ is the number of objects in w_l . The total number of objects is $n = \sum_{k=1}^K \sum_{l=1}^K n_{kl}$.

Table 2. Illustration of a confusion matrix

Verity decision	$w_1 \dots w_l \dots w_K$
w_1	$n_{11} \dots n_{1K}$
\vdots	\ddots
w_k	$\vdots \quad n_{kl} \quad \vdots$
\vdots	\ddots
w_K	$n_{K1} \dots n_{KK}$

3.5 Estimation of Contextual Relevance of a Source from the Confusion Matrix

In [7], Elouedi et al. define two ways to calculate the contextual relevance for each context: using the percentage of correct classifications for each context, and using a distance to determine the reliability rate of a source in the different contexts. The contextual weakening of a mass function based on its degree of contextual relevance is interesting. In our case, the application of this approach amounts to disjunctively combining the initial mass function with the 72 contextual elements that can be highlighted by the confusion matrix. This is very complicated from a computational point of view since it involves making 72 combinations for each mass function and for each source of information.

So, in the following we will use the confusion matrix to estimate the sincerity of the source.

3.6 Estimation of a Source Sincerity from the Confusion Matrix

It is possible to estimate the sincerity of a source always using a confusion matrix. In [17], Pichon et al. suggest the three approaches that can be used: generation of a Bayesian mass function characterizing the sincerity of a source, generation of a mass function from the pignistic probability, or generation of a mass function to characterize the sincerity of a source based on the work of Dempster [3].

As we mentioned, we are interested in the weakening of the mass function as proposed in Eq. 10. Indeed, in some applications, it is easy to define the set \bar{A} complement of A. This is even simpler when it comes to a small discernment space or where relevant knowledge is available (such as expert opinion). In our case, we can assimilate the set \bar{A} complement of A to the set of all the classes belonging to the discernment space Ω but not belonging to A. Assume that A is formed of 10 species, a mass will be transferred to the complement set \bar{A} composed of 62 species. The specificity of the information can be lost. For all these reasons, and considering the quality of the processed data as well as the available information, we were interested in defining the \bar{A} complement set for each context.

4 The Proposed Correction Method

The proposed correction method consists in reinforcing, contextually, the mass attributed to some classes. This correction requires retrieving information from the test database’s confusion matrix to define the complement set \overline{A}_i of each context c_i . The proposed approach consists of the following 3 steps:

- **construction of the Bayesian mass functions** from the probability distributions obtained at the output of a classifier
- **extraction of contextual meta-knowledges** from the confusion matrix. These meta-knowledges allow to define a complement set \overline{A}_i for each context c_i
- **contextual use of meta-knowledge** to correct the Bayesian mass functions

4.1 Building Bayesian Mass Functions

At the output of a classifier, we obtain a probability distribution $P_i(e_k)$ between the species. To move from probability theory to belief function theory, we construct Bayesian mass functions and transfer all available information to it as follows:

$$m_i\{e_k\} = P_i(e_k), \quad \forall e_k \in \Omega \tag{9}$$

$$m_i\{X\} = 0, \quad \text{if } |X| \neq 1 \tag{10}$$

where X is any set containing more than one species. Mass m_i equal to the probability P_i is given to each element e_k of Ω having a non-zero probability. A zero mass is given to all subsets.

4.2 Meta-knowledge Extraction

The purpose of this step is to use the confusion matrix to define, for each context c_i , the complement set \overline{A}_i . In our case, there are 72 contexts corresponding to the 72 species forming the discernment space Ω . In a confusion matrix, each line k represents the decision made in favor of w_k . Each column l represents the truth w_l . By traversing the matrix of confusion line by line without considering the diagonal, it is possible to extract, for a decision w_k , all the species in which the truth is found. This set is the complement set \overline{A}_k for a context c_k corresponding to w_k decision.

Example 2: Let’s consider the confusion matrix shown in the Table 3.

Let’s analyze the confusion matrix line by line without taking into account the diagonal. The first line corresponds to the decision e_1 . We notice through the third column that the classifier decides e_1 whereas it is e_3 . We thus consider, that for the context c_1 corresponding to the decision e_1 , the complement set is $\overline{A}_1 = \{e_3\}$. So, the idea is to assign a mass to e_3 each time the classifier decides e_1 .

Table 3. Example of confusion matrix

Truth decision	e_1	e_2	e_3
e_1	3	0	2
e_2	1	4	2
e_3	1	0	3

The second line corresponds to the decision e_2 . We notice through the first column that once on 5 the classifier decides e_2 when it is e_1 and through the third column that the classifier decides e_2 when it is e_3 . So a mass will be given to e_1 and e_3 each time the classifier decides e_2 .

The third line corresponds to the decision e_3 . According to the first column, once out of 5 the classifier decides e_3 whereas it is e_1 .

By going through the confusion matrix line by line, and following the classifier's decision, the three complementary sets corresponding respectively to a decision in favor of e_1 , e_2 and e_3 are obtained:

$$\begin{aligned}
 L(\text{decision} = e_1) &= L_1 = \overline{A_1} = \{e_3\} \\
 L(\text{decision} = e_2) &= L_2 = \overline{A_2} = \{e_1, e_3\} \\
 L(\text{decision} = e_3) &= L_3 = \overline{A_3} = \{e_1\}
 \end{aligned}
 \tag{11}$$

4.3 Using Meta-knowledge for Bayesian Mass Function Correction

Following the extraction of the complementary sets of each context, the correction of the mass functions is carried out by applying the approach presented in the Eq. 10. Applying this equation, must not change the Bayesian form of the mass function so that we can then build consonant mass functions. The information available will therefore be represented in the best possible way. The initial mass function will be weakened by β and a mass equal to $(1 - \beta)/|\overline{A_i}|$ is given to each element of the complement set. Depending on how the degree of sincerity β is calculated, the correction can be done in two ways:

1. β = overall degree of sincerity of the information source.
2. β = degree of contextual sincerity. In this case, the value of β is extracted from the diagonal of the confusion matrix for the 72 assumptions.

To weaken a mass function, the idea is to contextually vary the value of β . Thus, if the classifier decides that it is the species e_2 , the contextual value of β is that corresponding to the degree of sincerity of the classifier relative to the species e_2 . It is important to note that in our case, we consider that the decision of a classifier corresponds to the species with the maximum mass. Note that for most species, the context value of β is zero. Thus, in these cases, the initial mass function is multiplied by zero and a null mass is given to the complement set. In other words, the mass function becomes zero.

5 Experiments and Results: Application of the Mass Function Correction Method

5.1 Dataset Description

In the data set, there is 5104 leaves photos unequally distributed on 72 species. In fact, we have some species that are well presented in the dataset. However, we have some other species that are less presented and some species that have only two exemplary. We used 2572 leaves photos for training and the rest for testing.

5.2 What About Degree of Sincerity: Global or Contextual?

When the degree of sincerity is global, the value of β is the same for all contexts, while when the degree of sincerity is contextual, the value of β will depend on the context. We notice that in many contexts the value of β is null: 21 contexts for the Polygonal model classifier S_1 , 22 contexts for the Margin classifier S_2 and 34 contexts for the Base and Apex classifier S_3 . Even though contextual values of β usually allow contextual truthfulness to be taken into account in each case, in our case this approach is useless since, in most contexts, β have a null value.

5.3 Impact of the Proposed Method on the Classifier Performances

In order to evaluate the performance of the classifiers before and after the correction of the mass functions, the idea is to use the 10-fold cross-validation procedure. Meta-knowledge is extracted from 9/10th samples, and the effectiveness of the correction method is tested on the remaining 10. This operation is repeated 10 times, and each time the configuration and test samples are randomly drawn. The end result is the average value of the 10 results corresponding to the 10 repetitions. Figure 2 shows the results obtained before and after the correction of the mass functions of the test database. Since botanists like to see similar species, we have presented the results of the good classification rate according to the number N of species selected. We note a significant improvement in classifier performance for both modalities. Nevertheless, through Table 4 representing the

Table 4. Standard deviation of the different classifiers without and with the correction of mass functions, N is the number of returned species

Modality	Characteristics	Standard deviation					
		Without correction			With correction		
		N = 1	N = 5	N = 10	N = 1	N = 5	N = 10
Leaf	S_1 : Polygonal model	4.4604	4.7424	4.3574	4.4326	6.0894	2.9342
	S_2 : Margin	3.7012	4.3951	5.0837	5.3637	6.5911	4.4604
	S_3 : Base and Apex	5.4614	3.9912	3.2008	3.8133	3.8328	3.2924

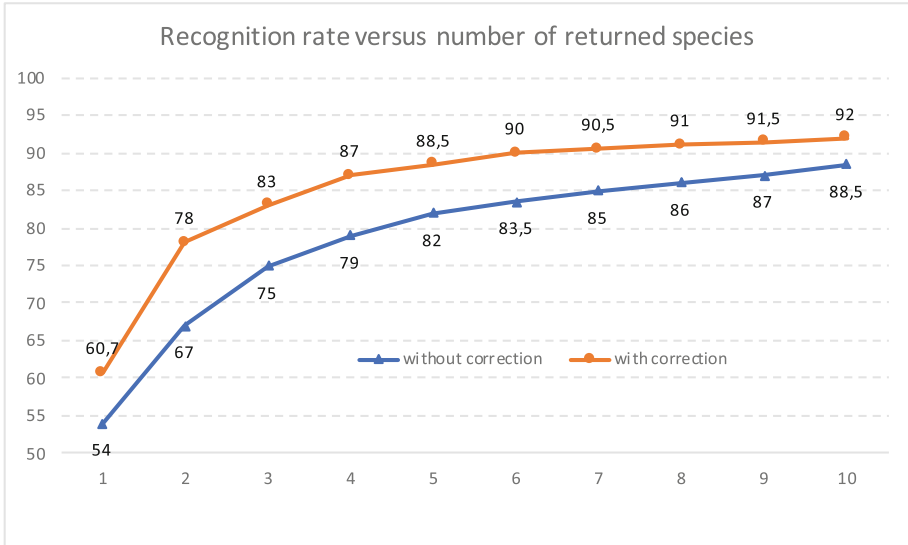


Fig. 2. Recognition rate without and with mass correction.

standard deviation for each classifier, we notice that the standard deviation is often important. Indeed, by randomly pulling the configuration data, the sets formed can be quite representative of the variations existing in the data as they can be weakly representative of these variations. The quality of the extracted meta-knowledge, therefore, varies according to the relevance of the configuration database, which subsequently influences the ability of the fusion system to generalize.

6 Conclusion

In this paper, we presented the different mechanisms for correcting mass functions and applied the most appropriate ones for our application. We have shown that the use of the confusion matrix allows us to define meta-knowledge useful for the correction of mass functions. We applied this technique to our merging system and showed that it significantly improved the rate of good classification. Work is in progress on the addition of the bark modality from which we can extract characteristics such as texture, color, ... which would allow us to differentiate between neighboring species. The bark-derived characteristics could serve as a meta-knowledge that we will use to resolve ambiguities when there are significant conflicts on species associated with leaf recognition. This would improve the performance of this system.

Acknowledgements. This work is part of ReVeRIES project supported by the French National Agency for Research with the reference ANR-15-CE38-0004.

References

1. Ameer, R.B., Valet, L., Coquin, D.: Influence of basic belief assignments construction on the behaviour of a fusion system for tree species recognition. In: 20th International Conference on Information Fusion, pp. 1–8. IEEE Fusion (2017)
2. Aregui, A., Denœux, T.: Constructing consonant belief functions from sample data using confidence sets of pignistic probabilities. *Int. J. Approximate Reasoning* **49**(3), 575–594 (2008)
3. Dempster, A.P.: New methods for reasoning towards posterior distributions based on sample data. *Ann. Math. Stat.* **37**, 355–374 (1966)
4. Denœux, T.: Inner and outer approximation of belief structures using a hierarchical clustering approach. *Int. J. Uncertainty Fuzziness Knowl. Based Syst.* **9**(04), 437–460 (2001)
5. Denœux, T., Smets, P.: Classification using belief functions: relationship between case-based and model-based approaches. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **36**(6), 1395–1406 (2006)
6. Dubois, D., Prade, H.: Representation and combination of uncertainty with belief functions and possibility measures. *Comput. Intell.* **4**(3), 244–264 (1988)
7. Elouedi, Z., Lefevre, E., Mercier, D.: Discountings of a belief function using a confusion matrix. In: 22nd IEEE International Conference on Tools with Artificial Intelligence (ICTAI), vol. 1, pp. 287–294. IEEE (2010)
8. Florea, M.C., Jousselme, A.L., Bossé, É., Grenier, D.: Robust combination rules for evidence theory. *Inf. Fusion* **10**(2), 183–197 (2009)
9. Kennes, R.: Computational aspects of the mobius transformation of graphs. *IEEE Trans. Syst. Man Cybern.* **22**(2), 201–223 (1992)
10. Lefevre, E., Colot, O., Vannoorenberghe, P.: Belief function combination and conflict management. *Inf. fusion* **3**(2), 149–162 (2002)
11. Lowrance, J.D., Garvey, T.D., Strat, T.M.: A framework for evidential-reasoning systems. In: Yager, R.R., Liu, L. (eds.) *Classic works of the Dempster-Shafer theory of belief functions. Studies in Fuzziness and Soft Computing*, pp. 419–434. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-44792-4_16
12. Martin, A., Osswald, C.: A new generalization of the proportional conflict redistribution rule stable in terms of decision. In: *Advances and Applications of DSMT for Information Fusion: Collected Works*, vol. 2, pp. 69–88 (2006)
13. Martin, A., Osswald, C.: Toward a combination rule to deal with partial conflict and specificity in belief functions theory. In: 10th International Conference on Information Fusion, pp. 1–8. IEEE (2007)
14. Mercier, D., Lefèvre, É., Delmotte, F.: Belief functions contextual discounting and canonical decompositions. *Int. J. Approximate Reasoning* **53**(2), 146–158 (2012)
15. Mercier, D., Quost, B., Denœux, T.: Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Inf. Fusion* **9**(2), 246–258 (2008)
16. Pichon, F., Dubois, D., Denœux, T.: Relevance and truthfulness in information correction and fusion. *Int. J. Approximate Reasoning* **53**(2), 159–175 (2012)
17. Pichon, F., Mercier, D., Lefevre, É., Delmotte, F.: Proposition and learning of some belief function contextual correction mechanisms. *Int. J. Approximate Reasoning* **72**, 4–42 (2016)
18. Shafer, G.: *A Mathematical Theory of Evidence*, vol. 1. Princeton University Press, Princeton (1976)

19. Smets, P.: Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *Int. J. Approximate Reasoning* **9**(1), 1–35 (1993)
20. Smets, P., Kennes, R.: The transferable belief model. *Artif. Intell.* **66**(2), 191–234 (1994)
21. Tessem, B.: Approximations for efficient computation in the theory of evidence. *Artif. Intell.* **61**(2), 315–329 (1993)