Yousri Kessentini, Thierry Paquet, Thomas Burger

Comparaison des méthodes probabilistes et évidentielles de fusion de classifieurs pour la reconnaissance de mots manuscrits

Yousri Kessentini* — **Thierry Paquet*** — **Thomas Burger****

* *Université de Rouen, Laboratoire LITIS EA 4108, site du Madrillet, St Etienne du Rouvray, France*

{yousri.kessentini,thierry.paquet}@univ-rouen.fr

** *Université Européenne de Bretagne, Université de Bretagne-Sud, CNRS, Lab-STICC, Centre de Recherche Yves Coppens, BP 573, F-56017 Vannes cedex, France*

thomas.burger@univ-ubs.fr

RÉSUMÉ. *Dans le cadre de la reconnaissance de mots manuscrits, nous proposons de comparer l'efficacité des méthodes probabilistes et évidentielles de combinaison de classifieurs MMC. Les performances obtenues indiquent que dans le cas de bases de données simples, les méthodes probabilistes sont plus efficaces. A l'inverse, sur des bases de données plus difficiles (lexique large, classifieurs individuels peu performants, etc.), la théorie des fonctions de croyance est plus efficace tout en ouvrant de nouvelles perspectives en termes de prise de décisions partielles.*

ABSTRACT. *Considering handwriting recognition, we compare the accuracy of probabilistic and eviential methods for ensemble HMM classifier combination. The recognition performances show that, in case of simple database, the probabilistic methods are more efficient. On the other hand, for more difficult recognition tasks (large vocabulary, weak classifiers, etc.), the belief functions theory provides better results, in addition to new insights for partial decision making.*

MOTS-CLÉS : *Combinaison de classfieurs, fonctions de croyance, modèle de croyance transferable, reconnaissance de l'écriture manuscrite.*

KEYWORDS: *Classifiers combination, belief functions, transferable belief model, handwriting recognition.*

1. Introduction

Depuis le début des années 90, la combinaison de classifieurs est une des directions les plus prometteuses pour augmenter les performances des systèmes de reconnaissance de formes (Xu *et al.,* 1992, Prevost *et al.,* 2003, Kim *et al.,* 2000). D'une manière générale, les informations issues des différents classifieurs peuvent être **imprécises** (pas assez focalisées), **incertaines** (admettant une composante aléatoire) ou **incomplètes** (ayant un point de vue partiel). Il arrive aussi, lorsque les informations manipulées proviennent de différentes sources, qu'elles soient en **conflit** : par exemple, il est possible qu'une source favorise l'hypothèse A, tandis qu'une autre favorise une hypothèse B différente de A ; ou encore qu'une relation modélisant une préférence n'implique pas un ordre complet sur l'ensemble des hypothèses.

La plupart des classifieurs efficaces sont formalisés d'un point de vue probabiliste (Duda *et al.,* 2001). Cependant, la théorie des probabilités ne propose pas d'outils spécifiques permettant de manipuler facilement ces notions d'imprécision, de conflit ou d'incomplétude. Ainsi, il y a deux manières d'aborder la combinaison de classifieurs probabilistes. La première consiste à convertir la sortie des classifieurs dans un autre formalisme permettant de gérer certaines de ces notions. Il existe plusieurs formalisme de ce genre : la théorie les fonctions de croyance (Shafer, 1976, Smets, 1994), la logique floue, la théorie de la résonance adaptative (Arbib, 2003), la théorie des possibilités (Dubois *et al.,* 2001), etc. La seconde manière consiste à conserver le formalisme probabiliste et à convertir l'imprécision, l'incomplétude ou le conflit en une incertitude supplémentaire. Ainsi, la logique probabiliste se détache de son origine exclusivement statistique afin de modéliser des connaissances subjectives, comme cela apparaît en inférence Bayésienne (Arbib, 2003, Duda *et al.,* 2001).

L'objectif de ce travail est de comparer ces deux approches dans le cas de la reconnaissance de l'écriture manuscrite multi-script. Lorsque le formalisme probabiliste n'est pas utilisé jusqu'au bout, nous faisons le choix de convertir la sortie de classifieurs dans le formalisme des fonctions de croyance. Ce choix est en partie justifié par la popularité grandissante de cette théorie, mais surtout par des travaux plus anciens le justifiant (Xu *et al.,* 1992).

Dans le domaine de la reconnaissance de formes, de très nombreux travaux ont déjà exploité la théorie des fonctions de croyances pour combiner les décisions de plusieurs classifieurs. Nous n'en citons donc que quelques uns : dans (Valente *et al.,* 2007), les auteurs présentent une méthode de combinaison de classifieurs de type réseaux de neurones pour la reconnaissance automatique de la parole. Dans (Burger *et al.,* 2008, Aran *et al.,* 2009), différents types de classifieurs (SVM et HMM respectivement) sont combinés pour la reconnaissance automatique de langages gestuels (la Langue Française Parlée Complétée et la Langue des Signes Américaine respectivement). Dans (Gagnon *et al.,* 2008), les auteurs fusionnent des informations audio et visuelles dans le cadre de la reconnaissance automatique de la parole. Dans (Mercier *et al.,* 2007), les auteurs présentent un modèle de fusion de décisions de lecteurs optiques afin d'améliorer la reconnaissance automatique d'adresses postales. Enfin, dans

(Masson *et al.,* 2009), les auteurs proposent un modèle théorique générique pour la combinaison de classifieurs dont les résultats sont exprimés sous forme évidentielle.

Cependant, à notre connaissance, il n'y a pas beaucoup de travaux exploitant la théorie des fonctions de croyances pour améliorer la reconnaissance de mots manuscrits. Xu, Krzyżak et Suen (Xu *et al.,* 1992) ont proposé une étude comparative des méthodes de combinaison de classifieurs appliquées à la reconnaissance de séquences numériques. Malgré des résultats équivalents pour de nombreuses méthodes, l'article conclut à une supériorité des combinaisons basées sur les fonctions de croyances, notamment en raison d'une plus grande robustesse quand les conditions se dégradent.

Ainsi, l'objectif de ce travail est triple : il s'agit d'abord d'appliquer la théorie des fonctions de croyance à la combinaison de classifieurs probabilistes pour la reconnaissance de l'écriture manuscrite, dans un contexte plus large que celui proposé dans (Xu *et al.,* 1992). Ensuite, il s'agit de comparer cela aux approches probabilistes en termes de taux de reconnaissance. Enfin, il s'agit de suivre les recommandations de (Xu *et al.,* 1992) et de voir ce que peut apporter cette théorie afin d'améliorer encore la modélisation du problème et donc d'améliorer la robustesse de la reconnaissance.

L'organisation de cet article est donc la suivante. Dans la section 2, les concepts de base des fonctions de croyance sont présentés. La section 3 présente les trois étapes permettant de convertir les sorties probabilistes des classifieurs en des fonctions de croyance et de prendre une décision. Ces étapes sont (1) le choix du cadre de discernement, (2) la génération des fonctions de croyance, et (3) la combinaison des fonctions de croyance suivie de la prise de décision. La section 4 présente les résultats expérimentaux en reconnaissance de mots manuscrits latins et arabes. Enfin, la section 5 discute des résultats et évoque les possibilités supplémentaires qu'offre la théorie des fonctions de croyance.

2. Le Modèle des Croyances Transférables

La théorie des fonctions de croyance (ou théorie de Dempster-Shafer) a été formalisée par G. Shafer (Shafer, 1976) à partir de travaux antérieurs du statisticien A. Dempster. Par la suite, Ph. Smets (Smets, 1993) fut l'un des principaux contributeurs à cette théorie grâce à son **Modèle des Croyances Transférables** (MCT), qui propose une axiomatique des fonctions de croyance indépendante de celle de la théorie des probabilités.

Les mécanismes de raisonnement du MCT peuvent être regroupés en deux niveaux : le **niveau crédal** permet la représentation des informations par des fonctions de croyance, ainsi que la combinaison de ces dernières afin de modéliser la fusion des informations correspondantes. Le **niveau pignistique** permet la prise de décision dans un cadre probabiliste facile à interpréter dans lequel on retrouve la notion d'hypothèse et de pari statistiquement gagnant.

2.1. *Niveau crédal*

Soit $\Omega = \{\omega_1, ..., \omega_K\}$ un ensemble fini, appelé **cadre de discernement** et représentant des hypothèses exhaustives et exclusives. Une **fonction de croyance** bel est une fonction de l'**ensemble des parties** (ou **powerset**) $\mathcal{P}(\Omega)$ dans $[0, 1]$ définie par :

$$bel\,(A) = \sum_{B \subseteq A, B \neq \emptyset} m\,(B)\,, \qquad \forall A \subseteq \Omega \qquad [1]$$

où $m : \mathcal{P}(\Omega) \mapsto [0, 1]$, appelé **fonction de masses**, vérifie : $\sum_{A \subseteq \Omega} m\,(A) = 1$

Un sous-ensemble $A \subseteq \Omega$ tel que $m\,(A) > 0$ est appelé **élément focal** de m. La masse $m(A)$ représente le degré de croyance attribué à la proposition A et qui n'a pas pu, compte-tenu de l'état de la connaissance, être affectée à un sous-ensemble plus spécifique que A. A l'inverse, $bel(A)$ correspond à la croyance en toutes les hypothèses qui impliquent A. Notons que bel et m encodent la même information, et qu'il est possible de passer de manière biunivoque de l'une à l'autre. Il est possible de représenter cette information avec encore d'autres structures. Les plus utilisées sont les suivantes : la **plausibilité** pl d'un élément A est la part maximale de croyance qui pourrait soutenir A :

$$pl\,(A) = \sum_{B \cap A \neq \emptyset} m\,(B)\,, \forall A \subseteq \Omega \qquad [2]$$

La **communalité** q d'un élément A représente la somme des masses allouées aux sur-ensembles de A et donc qui ont A en commun :

$$q\,(A) = \sum_{A \subseteq B} m\,(B)\,, \forall A \subseteq \Omega \qquad [3]$$

Deux fonctions de croyance, bel_1 et bel_2, issues de sources d'information indépendantes et fiables peuvent être combinées en une fonction $bel_\cap = bel_1 \bigcirc\!\!\!\!\cap\, bel_2$ en utilisant la **combinaison de Dempster**. Celle-ci est définie à partir des masses de croyance :

$$\forall A \subseteq \Omega \qquad m_\cap(A) = \frac{1}{1 - \mathcal{K}_{12}} \sum_{B \cap C = A} m_1\,(B) \cdot m_2\,(C) \qquad [4]$$

$$\text{avec} \qquad \mathcal{K}_{12} = \sum_{B \cap C = \emptyset} m_1\,(B) \cdot m_2\,(C) \qquad [5]$$

où $\mathcal{K}_{12}$ est une mesure du conflit entre les sources d'information 1 et 2. Notons enfin que la combinaison de Dempster se calcule aussi très simplement à partir des fonctions de communalités : $q_\cap(A) = q_1\,(A) \cdot q_2\,(A)\,,\ \forall A \subseteq \Omega$.

2.2. *Niveau pignistique*

Dans le MCT, Smets propose de prendre une décision non pas sur une des structures du niveau crédal (croyance, plausibilité, etc.) mais sur une distribution de probabilités. Il définit donc la **probabilité pignistique** $BetP$ sur Ω de la manière suivante :

$$BetP\,(\omega_i) = \frac{1}{1 - m(\emptyset)} \sum_{A \ni \omega_i} \frac{m\,(A)}{|A|} \qquad \forall \omega_i \in \Omega \tag{6}$$

où $|A|$ représente la cardinalité du sous-ensemble $A \subseteq \Omega$. Cette transformation signifie simplement que la croyance en un élément non-singleton A doit être également distribuée entre les hypothèses singletons qui la constituent puisqu'il n'y a aucune raison de favoriser l'une plutôt que les autres : c'est le principe de la raison insuffisante (Keynes, 1921). Une fois $BetP$ calculée, la décision est finalement prise en choisissant l'élément ω_* possédant la plus grande probabilité pignistique.

3. Méthode de reconnaissance proposée

Le problème fondamental en reconnaissance d'écriture manuscrite est de trouver, à partir d'une séquence d'observations $O = (o_1, o_2, ..., o_T)$ de longueur T fournie en entrée, et d'un vocabulaire (lexique) R_V de V mots différents, le mot $\omega_* \in R_V$ qui correspond le mieux à O.

Les modèles de Markov cachés (MMC) sont un des outils permettant de résoudre ce problème en évaluant une vraisemblance $L(O|\omega_k)$ pour chaque modèle de mot $\omega_k,\ k \leq V$ du lexique[1]. La sortie du classifieur est la liste des N meilleures hypothèses de mots ainsi que leurs vraisemblances.

Dans les expérimentations menées dans ce travail, nous utilisons trois sources d'informations différentes : des contours supérieurs, des contours inférieurs, et des densités de pixels (Kessentini *et al.,* 2010). Un classifieur de type MMC est utilisé pour chaque source d'information. Il s'agit donc ensuite de combiner les sorties de ces classifieurs pour prendre une décision. Nous décrivons donc dans la suite les étapes de cette combinaison dans le cadre du MCT.

3.1. *Choix du cadre de discernement*

Dans un premier temps, il est nécessaire de définir le cadre de discernement, c'est-à-dire dans notre application, l'ensemble des mots sur lequel sont définies les 3 fonctions de croyance récupérées en sortie des classifieurs. Le plus simple est de considérer comme cadre de discernement Ω l'ensemble de tous les mots du lexique. Le problème

1. Théoriquement, cette vraisemblance se calcule via l'algorithme Forward, cependant en pratique, l'algorithme de Viterbi fournit une approximation souvent utilisée.

dans ce cas est que si nous travaillons avec un lexique de V mots, il faudra spécifier $|\mathcal{P}(\Omega)| = 2^V$ valeurs pour chaque fonction de croyance. Cette explosion combinatoire empêchera de généraliser l'approche sur de grands lexiques.

Pour contourner cette difficulté, nous avons choisi de ne considérer comme éléments du cadre de discernement que les K mots communs retournés par les différents classifieurs (Rang N, avec $N \geq K$). Ainsi nous pouvons contrôler la taille des supports de nos différentes fonctions de croyance qui sera toujours inférieure à 2^N. Notons que plus le nombre de classifieurs à combiner est important, plus K, le nombre de mots communs à tous les classifieurs, risque de diminuer. Ainsi, afin d'éviter que le cadre de discernement ne soit trop petit, il peut être judicieux d'augmenter le nombre N d'hypothèses considérées. Bien sûr, le cadre gardant le même ordre de grandeur, la complexité calculatoire n'en pâtit pas. Notons que d'un point de vue conceptuel, cette méthode est proche de celle proposée par Tessem (Tessem, 1993), afin de limiter le nombre d'éléments focaux.

3.2. *Construction des fonctions de masse*

Ayant fixé le cadre de discernement Ω, l'étape suivante consiste à construire la fonction de masse associée à chaque classifieur. Plusieurs méthodes de construction de modèles de fonctions de croyance à partir d'un ensemble d'apprentissage ont été proposées dans la littérature (Vannoorenberghe, 2003). Dans notre cas, nous nous intéressons à leurs définition à partir d'un ensemble de vraisemblances. En pratique, de nombreux classifieurs probabilistes ne fournissent pas une vraisemblance, mais une log-vraisemblance, ou encore une vraisemblance ré-echelonnée à divers instants du processus de calcul, afin d'éviter les débordements de mémoire sur machine.

Ainsi, selon les cas, il s'agit de convertir un potentiel (Shafer *et al.,* 1988) à valeurs dans $\mathbb{R}$, à valeurs dans $\mathbb{R}^+$, où à valeurs dans $[0, 1]$ en une fonction de masse (et de répéter cette opération pour chaque classifieur).

Le cas le plus simple est celui où les valeurs appartiennent à $[0, 1]$. Il y a deux manières de traiter ce cas, que nous présentons tour à tour.

Transformation consonante

La première méthode de conversion (Dubois *et al.,* 2001) consiste à normaliser les vraisemblances issues des classifieurs afin de les interpréter comme une distribution de probabilités subjectives. Ensuite, on convertit cette distribution en une **fonction de croyance consonante**. Un jeu de masses consonantes est tel que les éléments focaux $\mathcal{F} = \{F_1, F_2, \ldots, F_c\}$ sont ordonnés par rapport à l'inclusion : $\exists$ une permutation σ telle que $F_{\sigma(1)} \subset F_{\sigma(2)} \subset \ldots \subset F_{\sigma(c)}$, ce qui implique qu'il y aie au plus $|\Omega| = K$ éléments focaux. On a donc $c \leq K$.

Soit p une fonction de probabilité subjective déduite des vraisemblances par normalisation afin que leur somme soit égale à 1. Naturellement, p est définie sur Ω.

Par souci de simplicité de notations, nous supposons que les éventualités de Ω, $\{\omega_1, \omega_2, \ldots, \omega_K\}$ sont déjà ordonnées par valeurs de probabilité p décroissantes : $p(\omega_1) \geq p(\omega_2) \geq \ldots \geq p(\omega_K)$. Soit m la fonction de croyance consonante correspondant à p. On a :

$$\begin{aligned} m\left(\{\omega_1, \omega_2, \ldots, \omega_K\}\right) = m\left(\Omega\right) &= K \times p(\omega_K) \\ m\left(\{\omega_1, \omega_2, \ldots, \omega_k\}\right) &= k \times [p(\omega_k) - p(\omega_{k+1})] \qquad \forall\, k < K \\ m\left(.\right) &= 0 \qquad \text{sinon} \end{aligned}$$

La règle dérivée du GBT

Le GBT (Théorème de Bayes Généralisé) est la généralisation du théorème de Bayes aux fonctions de croyance (Smets, 1993). Dans la littérature, il est convenu que la plausibilité de ω_k sachant O est assimilée à une vraisemblance $L\left(O|\omega_k\right)$ (Smets, 1993). Étant donné la vraisemblance $L\left(O|\omega_k\right)$ pour chaque $\omega_k \in \Omega$, on en déduit donc une construction possible des fonctions de croyance à partir des vraisemblances pour chaque $A \subseteq \Omega$.

$$m\left(A\right) = \prod_{\omega_k \in A} L\left(O|\omega_k\right) \prod_{\omega_k \in \overline{A}} \left(1 - L\left(O|\omega_k\right)\right) \qquad [7]$$

où encore à partir des communalités :

$$q\left(A\right) = \prod_{\omega_k \in A} L\left(O|\omega_k\right) \qquad [8]$$

Dans le cas continu, les vraisemblances ne sont pas comprises dans $[0, 1]$, mais dans $\mathbb{R}^+$. Dès lors, on peut utiliser une méthode classique proposée dans (Appriou, 1991).

La règle d'Appriou

A partir d'une fonction de vraisemblance $L\left(O|\omega_k\right)$ définie dans $[0, +\infty]$, il est possible de définir une fonction de masse m_k associée à la classe ω_k. Les éléments focaux de cette fonction sont $\overline{\omega_k} = \{\Omega \setminus \omega_k\}$ le complémentaire du singleton $\{\omega_k\}$ et l'ensemble Ω lui-même. Cette fonction de masse est définie de la manière suivante :

$$\begin{aligned} m_k\left(\{\omega_k\}\right) &= 0 & [9] \\ m_k\left(\{\overline{\omega_k}\}\right) &= \alpha_k \left(1 - R.L\left(O|\omega_k\right)\right) & [10] \\ m_k\left(\Omega\right) &= 1 - \alpha_k \left(1 - R.L\left(O|\omega_k\right)\right) & [11] \end{aligned}$$

Dans ces équations, α_k est un coefficient d'affaiblissement associé à la classe ω_k tandis que R est un coefficient de normalisation positif inférieur ou égal à

$\left[\max_k \left(L\left(O|\omega_k\right)\right)\right]^{-1}$. A partir de ces K fonctions de croyance et en utilisant la règle de combinaison conjonctive, une fonction de masse unique m peut être obtenue :

$$m = m_1 \textcircled{\cap} m_2 \textcircled{\cap} \ldots \textcircled{\cap} m_K$$

Une méthode similaire est utilisée dans (Burger *et al.,* 2008) afin de construire une fonction de masse à partir des sorties d'un banc de classifieurs à marges. Comme cela est montré dans (Burger *et al.,* 2008), cette méthode de construction permet aussi dans certains cas d'aboutir à la transformation consonante décrite plus haut.

Cas des log-vraisemblances

Dans le cas où, comme avec l'algorithme de Viterbi, des log-vraisemblances (donc comprises dans $[-\infty, +\infty]$) sont fournies en sortie des classifieurs, le plus simple est d'appliquer une fonction de $f : [-\infty, +\infty] \mapsto [0, 1]$ à l'ensemble des log-vraisemblances. Généralement, on choisie f dans la famille des fonctions sigmoïdes. Nous proposons, pour chaque distribution de log-vraisemblances fournie en sortie d'un classifieur, d'utiliser la fonction sigmoïde suivante :

$$f(x) = \frac{1}{1 + e^{-\lambda(x-\tilde{x})}} \quad \text{avec} \quad \lambda = \frac{1}{\max_i |L_i - \tilde{x}|} \qquad [12]$$

où $\tilde{x}$ est la médiane des valeurs des log-vraisemblances. Ainsi, la sigmoïde appliquée est centrée sur la médiane (moins sensible aux valeurs extrêmes que la moyenne) des vraisemblances, et avec λ défini par rapport à la vraisemblance la plus "éloignée" de la médiane.

Ensuite, ces vraisemblances sont transformées en une fonction de masse en utilisant les transformations précédemment décrites.

3.3. *Combinaison et prise de décision*

A ce stade, le cadre de discernement est fixé, et une fonction de masses est construite à partir des vraisemblances issues de chacun des classifieurs MMC. Il ne reste donc plus qu'à les combiner en utilisant la règle de combinaison conjonctive (l'équation 4). A partir de l'unique fonction de masse résultante, le calcul de la probabilité pignistique peut se faire à l'aide de l'équation 6. Le mot reconnu est simplement celui qui permet de maximiser cette probabilité pignistique.

4. Expérimentation

Pour la phase d'expérimentation, nous avons choisi de travailler sur deux bases de mots publiques différentes, la base IRONOFF (Viard-Gaudin *et al.,* 1999) de mots Latins et la base IFN/ENIT (Pechwitz *et al.,* 2002) de mots Arabes. Ce choix nous permet de confronter nos algorithmes et leurs paramétrages à différentes tailles de lexique, et à différents scripts.

Les trois méthodes de transformations vraisemblance-masse présentées sont testées : Consonante, GBT, Appriou. Pour les transformations consonante et dérivée du GBT, la log-vraisemblance est re-échelonnée sur [0,1] avec la fonction sigmoïde décrite plus haut. Pour la règle d'Appriou, nous avons fait de même en considérant simplement que $[0, 1] \subseteq [-\infty, +\infty]$; ensuite, de manière classique (Mathevet *et al.,* 1999), nous avons choisi $\alpha = 0.9$ et $R = \max_i(L_i)$.

Ces trois méthodes sont confrontées à trois méthodes probabilistes classiques : la somme des vraisemblances en sorties des classifieurs, le produit des vraisemblances en sortie des classifieurs et le vote de Borda Count, qui consiste à associer un score à chaque hypothèse en fonction de son classement par chacun des classifieurs (Kittler *et al.,* 1998).

4.1. *Résultats sur la base IFN/ENIT*

La base de données de référence IFN-ENIT comporte 32492 mots arabes écrits par 411 scripteurs. Les mots de cette base correspondent à un lexique de 946 noms de villes et villages tunisiens. Cinq ensembles distincts de données (a, b, c, d, e) sont prédéfinies dans la base IFN/ENIT. Dans notre cas, les sous-ensembles a, b, c et d ont été utilisés pour l'apprentissage et le sous-ensemble e pour le test.

Le tableau 1 présente les résultats de reconnaissance des classifieurs MMC sur différentes sources d'informations, sous la forme du taux de réponses pour lesquels la bonne solution est dans le rang 1 ou 2 des hypothèses retenues.

MMC	Rang 1	Rang 2
1	73.60	79.77
2	65.90	74.03
3	72.97	79.73

Tableau 1. *Résultats de reconnaissance des différents classifieurs MMC sur la base IFN/ENIT (1 : Contour Supérieur, 2 : Contour inférieur, 3 : Densité).*

Le Tableau 2 présente les résultats de la combinaison des différents classifieurs MMC dans le cadre du MCT. Les trois règles de transformations vraisemblance-masse présentées dans la section 3.2 sont testées. Notons que ces résultats sont obtenus en ne considérant que les dix meilleures propositions de mots retournées par chaque classifieur (Rang 10) lors de la construction du cadre de discernement. Ces résultats montrent que les trois règles de construction des fonctions de masse donnent des résultats similaires, avec un petit avantage pour les règles d'Appriou et du GBT.

MMC	Consonante		GBT		Appriou	
	Rang 1	**Rang 2**	**Rang 1**	**Rang 2**	**Rang 1**	**Rang 2**
1-2	75.37	80.40	76.53	81.73	76.53	81.73
1-3	78.73	83.13	78.80	83.17	78.80	83.17
2-3	75.83	80.43	75.87	80.40	75.87	80.40
1-2-3	79.07	82.43	79.20	82.50	79.20	82.50

Tableau 2. *Résultats de la combinaison MCT sur la base IFN/ENIT en considérant les 10 meilleures hypothèses retournées par chaque classifieur pour la définition du cadre de discernement.*

L'influence du nombre de propositions de mots considérées par chaque classifieur lors de la construction du cadre de discernement a aussi été évaluée. Ainsi, les résultats représentés dans le tableau 3 correspondent au cas où ce nombre est de 20. Ces résultats montrent qu'on arrive à améliorer légèrement les performances par rapport aux résultats du tableau 2, pour lequel 10 propositions sont considérées. D'autres expériences ont été réalisées pour étudier l'influence de ce paramètre sur les performances du système de reconnaissance et ont montré que le fait de l'augmenter à plus de 20 n'améliore plus les performances, mais augmente la complexité calculatoire.

MMC	Consonante		GBT		Appriou	
	Rang 1	**Rang 2**	**Rang 1**	**Rang 2**	**Rang 1**	**Rang 2**
1-2	76.50	82.30	76.87	82.43	76.87	82.43
1-3	79.40	84.17	79.50	84.23	79.50	84.23
2-3	76.97	82.30	77.07	82.30	77.07	82.30
1-2-3	79.63	83.40	79.73	83.45	79.73	83.45

Tableau 3. *Résultats de la combinaison MCT sur la base IFN/ENIT en considérant les 20 meilleurs hypothèses retournées par chaque classifieur pour la définition du cadre de discernement.*

MMC	Somme		Produit		Vote	
	Rang 1	**Rang 2**	**Rang 1**	**Rang 2**	**Rang 1**	**Rang 2**
1-2	76.00	80.53	76.27	80.70	76.27	80.07
1-3	76.90	82.13	79.43	83.17	77.77	83.03
2-3	72.97	79.47	76.67	80.50	74.63	80.20
1-2-3	78.47	82.87	79.53	83.10	79.43	83.20

Tableau 4. *Résultats des méthodes de combinaison probabilistes sur la base IFN/ENIT*

De manière similaire, les résultats obtenus par les méthodes probabilistes sont donnés dans le tableau 4.
Nous remarquons que la combinaison basée sur le MCT donne des performances légèrement meilleures que les règles de combinaison probabilistes par somme, produit et vote de Borda Count, mais qu'elles sont globalement du même ordre.

4.2. *Résultats sur la base IRONOFF*

IRONOFF est une base de données duale en-ligne et hors-ligne. Le sous-ensemble IRONOFF-Chèque est restreint à un petit lexique d'environ 30 mots utilisés pour exprimer les montants de chèques français. On dispose de 7956 images de mots pour l'apprentissage et de 3987 images de mots pour le test. Comme précédemment, le tableau 5 présente les résultats de reconnaissance des classifieurs MMC correspondant aux différentes sources d'informations.

MMC	**Rang 1**	**Rang 2**
1	85.65	91.51
2	79.59	89.32
3	90.25	95.45

Tableau 5. *Résultats de reconnaissance des classifieurs MMC sur la base IRONOFF*

Le Tableau 6 présente les résultats de la combinaison des différents classifieurs MMC en utilisant la combinaison basée sur le MCT. Là encore, ces résultats sont obtenus en ne considérant que le Rang 10 de chaque classifieur lors de la construction du cadre de discernement. Les résultats montrent que les trois transformées vraisemblance-masse utilisées donnent des performances similaires conformément aux résultats obtenus sur la base de mots Arabe.

MMC	**Consonante**		**GBT**		**Appriou**	
	Rang 1	**Rang 2**	**Rang 1**	**Rang 2**	**Rang 1**	**Rang 2**
1-2	89.39	93.64	89.37	93.64	89.37	93.64
1-3	92.61	96.10	92.64	96.13	92.64	96.13
2-3	92.54	95.93	92.54	95.90	92.54	95.90
1-2-3	92.76	95.78	92.79	95.75	92.79	95.83

Tableau 6. *Résultats de la combinaison MCT sur la base IRONOFF en considérant les 10 meilleurs hypothèses retournées par chaque classifieur pour la définition du cadre de discernement.*

Si un Rang 15 est utilisé au lieu d'un Rang 10 lors de la définition du cadre de discernement, nous obtenons les résultats représentés dans le tableau 7. Comme précédemment sur l'autre jeu de données, ces résultats montrent qu'on arrive à améliorer légèrement les performances par rapport aux résultats du tableau 6 ; cependant, d'autres expériences ont montré que le fait d'augmenter ce nombre au-delà de 15 n'améliore pas les performances de classification de manière sensible, mais réduit l'efficacité calculatoire.

MMC	Consonante		GBT		Appriou	
	Rang 1	Rang 2	Rang 1	Rang 2	Rang 1	Rang 2
1-2	89.44	93.89	89.41	93.82	89.40	93.82
1-3	92.86	96.46	92.89	96.43	92.89	96.42
2-3	92.76	96.31	92.76	96.33	92.76	96.31
1-2-3	93.29	96.63	93.36	96.68	93.29	96.63

Tableau 7. *Résultats de la combinaison MCT sur la base IRONOFF en considérant les 15 meilleurs hypothèses retournées par chaque classifieur pour la définition du cadre de discernement.*

Les mêmes règles de combinaison probabilistes que précédemment donnent sur cette base les résultats du tableau 8. Nous remarquons que la combinaison basée sur le MCT donne des performances légèrement meilleures que les deux règles de combinaison probabilistes par somme et par vote de Borda Count. Cependant, les performances de la règle de combinaison par produit de vraisemblances sont légèrement meilleures.

MMC	Somme		Produit		Vote	
	Rang 1	Rang 2	Rang 1	Rang 2	Rang 1	Rang 2
1-2	89.47	94.52	89.62	94.22	85.15	92.81
1-3	90.85	94.90	93.82	97.16	90.63	95.35
2-3	88.09	94.04	93.37	96.83	87.81	94.75
1-2-3	91.68	95.70	94.07	97.54	91.15	95.58

Tableau 8. *Résultats des méthodes de combinaison probabilistes sur la base IRONOFF*

5. Discussion

Nous avons proposé dans ce travail une étude comparative de méthodes de fusion de classifieurs. Les méthodes évidentielles proposées (déclinées dans trois versions différentes) permettent de pallier l'un des défauts majeurs de la théorie des fonctions de croyance, à savoir l'explosion combinatoire résultant de la non maîtrise de la taille du cadre de discernement. Dans l'ensemble, les méthodes évidentielles donnent des résultats équivalents aux méthodes probabilistes, même si ces dernières sont très légèrement moins performantes, dans tous les cas étudiés à l'exception d'un seul. En effet, le produit des vraisemblances semble être plus efficace que toutes les méthodes évidentielles ou probabilistes, dans le cas de la base IRONOFF, pour laquelle les taux de reconnaissance des classifieurs individuels sont les plus élevés, et pour laquelle le lexique est plus petit que pour la base IFN/ENIT. Ainsi, les résultats que nous obtenons pour la combinaison de classifieurs probabilistes concordent avec les conclusions de (Xu *et al.,* 1992) sur les classifieurs de Type 1 (qui fournissent simplement l'index de la classe, et non une vraisemblance), indiquant qu'à performance égale, la robustesse est le principal avantage de ces méthodes.

Dans le cas d'une application ne nécessitant pas une telle robustesse, les méthodes probabilistes ont toutefois l'avantage de la simplicité d'implémentation et du coût calculatoire. Rappelons malgré tout qu'une définition judicieuse du cadre de discernement permet de limiter ce coût.

Malgré tout, la modélisation que nous proposons n'est pas exempte de critique. En effet, l'utilisation de la combinaison conjonctive suppose que les sources d'informations sont indépendantes, i.e. que les classifieurs sont indépendants, ce qui n'est évidemment pas le cas. Dans une optique simplificatrice pour notre modèle, nous avons relâché la contrainte d'indépendance, tout comme de nombreux travaux du même ordre. Une modélisation plus fine pourrait être envisagée, ou *a contrario*, des classifieurs plus indépendants pourraient être considérés (bases d'apprentissage différentes, algorithmes d'inférence différents, etc.). A ce titre, notons que notre méthode est facilement adaptable à tout autre classifieur pouvant fournir une réponse de type probabiliste. Ainsi, dans (Burger *et al.,* 2008), une méthode similaire est utilisée pour la combinaison de SVM.

Enfin, notons que l'un des principaux avantages des méthodes évidentielles est de permettre une modélisation plus riche. Ainsi, dans (Xu *et al.,* 1992), il est proposé de pondérer les masses de croyance en sortie de chacun des classifieurs en fonction du taux de reconnaissance de chacun d'eux pris individuellement. Il est aussi possible de tirer parti de la masse de conflit issue de la combinaison conjonctive, puisqu'elle peut être interprétée comme la probabilité que la bonne décision se retrouve en dehors du cadre. Enfin, des méthodes de décisions mixtes, telles que celles utilisées dans (Aran *et al.,* 2009) permettent de déterminer la taille N de la liste la plus adaptée à chaque décision, afin de limiter les erreurs. Quoi qu'il en soit, l'utilisation du formalisme évidentiel ouvre de nouvelles perspectives pour l'amélioration des taux de reconnaissance. Ceux-ci étant déjà relativement élevés, le gain en terme de point gagné semble toujours modeste, même quand la proportion d'erreurs restante ainsi évitée est conséquente.

6. Conclusion

Nous avons comparé dans cet article des méthodes de combinaison de classifieurs MMC pour la reconnaissance de mots manuscrits latins et arabes. Nous avons considéré d'un côté trois méthodes basées sur les fonctions de croyances et de l'autre côté des méthodes probabilistes. En dépit de résultats très proches, nos expériences permettent de conclure à un léger avantage des méthodes basées sur les fonctions de croyance, principalement pour les bases de données difficiles pour lesquelles des classifieurs individuels ne permettent pas de bons résultats. Ainsi, nous pensons que l'utilisation de ce formalisme, même s'il ne révolutionnera pas l'état de l'art, est à même de permettre d'améliorer les taux de reconnaissances par rapport aux méthodes probabilistes plus classiques.

7. Bibliographie

Appriou A., « Probabilités et incertitude en fusion de données multi-senseurs », *Revue Scientifique et Technique de la Défense*, vol. 11, p. 27-40, 1991.

Aran O., Burger T., Caplier A., Akarun L., « A Belief-Based Sequential Fusion Approach for Fusing Manual and Non-Manual Signs », *Pattern Recognition*, vol. 42, n° 5, p. 812-822, May, 2009.

Arbib M. A., *The Handbook of Brain Theory and Neural Networks*, Second Edition, MIT Press, 2003.

Burger T., Aran O., Urankar A., Akarun L., Caplier A., « A Dempster-Shafer Theory based combination of classifiers for hand gesture recognition », *Computer Vision and Computer Graphics - Theory and Applications, Lecture Notes in Communications in Computer and Information Science*, 2008.

Dubois D., Prade H., Smets P., « New Semantics for Quantitative Possibility Theory », *in* , S. Benferhat, , P. Besnard (eds), *Proc. of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU 2001)*, Springer-Verlag, Toulouse, France, p. 410-421, 2001.

Duda R., Hart P., Stork D., *Pattern Classification*, Wiley, 2001.

Gagnon L., Foucher S., Laliberte F., Boulianne G., « A simplified audiovisual fusion model with application to large-vocabulary recognition of French Canadian speech », *Electrical and Computer Engineering, Canadian Journal of*, vol. 33, n° 2, p. 109-119, Spring, 2008.

Kessentini Y., Paquet T., Hamadou A. B., « Off-line handwritten word recognition using multi-stream hidden Markov models », *Pattern Recognition Letters*, vol. 30, n° 1, p. 60-70, 2010.

Keynes J. M., « Fundamental Ideas », *A Treatise on Probability, Ch. 4*, 1921.

Kim J. H., Kim K. K., Nadal C. P., Suen C. Y., « A Methodology of Combining HMM and MLP Classifiers for Cursive Word Recognition », *International Conference on Pattern Recognition*, vol. 2, p. 319-322, 2000.

Kittler J., Hatef M., Duin R. P., Matas J., « On Combining Classifiers », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, n° 3, p. 226-239, 1998.

Masson M.-H., Denoeux T., « Belief Functions and Cluster Ensembles », *ECSQARU*, p. 323-334, July, 2009.

Mathevet S., Trassoudaine L., Checchin P., Auzon J., « Combinaison de segmentations en régions », *Traitement du signal*, 1999.

Mercier D., Cron G., Denoeux T., Masson M.-H., « Fusion de décisions postales dans le cadre du Modèle des Croyances Transférables », *Traitement du Signal*, vol. 24, n˚ 2, p. 133-151, 2007.

Pechwitz M., Maddouri S., Maergner V., Ellouze N., Amiri H., « IFN/ENIT - database of Handwritten Arabic Words », *Colloque International Francophone sur l'Ecrit et le Doucement*, vol. , p. 129-136, 2002.

Prevost L., Michel-Sendis C., Moises A., Oudot L., Milgram M., « Combining model-based and discriminative classifiers : application to handwritten character recognition », *International Conference on Document Analysis and Recognition*, vol. 1, p. 31-35, 2003.

Shafer G., *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

Shafer G., Shenoy P. P., « Local Computation on hypertrees », *Working paper No. 201, School of Business, University of Kansas*, 1988.

Smets P., « Belief functions : the disjunctive rule of combination and the generalized Bayesian theorem », *Int. Journal of Approximate Reasoning*, vol. 9, p. 1-35, 1993.

Smets P., « The transferable belief model », *Artif. Intell.*, vol. 66, n˚ 2, p. 191-234, 1994.

Tessem B., « Approximations for Efficient Computation in the Theory of Evidence », *Artif. Intell*, vol. 61, n˚ 2, p. 315-329, 1993.

Valente F., Hermansky H., « Combination of Acoustic Classifiers Based on Dempster-Shafer Theory of Evidence », *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, vol. 4, p. 1129-1132, April, 2007.

Vannoorenberghe P., « Un état de l'art sur les fonctions de croyance appliquées au traitement de l'information », *Information interaction intelligence*, vol. 3, n˚ 2, p. 9-45, 2003.

Viard-Gaudin C., Lallican P. M., Binter P., Knerr S., « The IRESTE On/Off (IRONOFF) Dual Handwriting Database », *International Conference on Document Analysis and Recognition*, vol. 0, p. 455-458, 1999.

Xu L., Krzyzak A., Suen C., « Methods of combining multiple classifiers and their applications to handwriting recognition », *IEEE Trans. Syst., Man, Cybern.*, 1992.

ANNEXE POUR LE SERVICE FABRICATION
A FOURNIR PAR LES AUTEURS AVEC UN EXEMPLAIRE PAPIER
DE LEUR ARTICLE ET LE COPYRIGHT SIGNE PAR COURRIER
LE FICHIER PDF CORRESPONDANT SERA ENVOYE PAR E-MAIL

1. ARTICLE POUR LES ACTES :

   Yousri Kessentini, Thierry Paquet, Thomas Burger

2. AUTEURS :

   Yousri Kessentini — Thierry Paquet* — Thomas Burger***

3. TITRE DE L'ARTICLE :

   Comparaison des méthodes probabilistes et évidentielles de fusion de classifieurs pour la reconnaissance de mots manuscrits

4. TITRE ABRÉGÉ POUR LE HAUT DE PAGE MOINS DE 40 SIGNES :

   Combinaison évidentielle de classifieurs

5. DATE DE CETTE VERSION :

   5 février 2010

6. COORDONNÉES DES AUTEURS :

   – adresse postale :
   * Université de Rouen, Laboratoire LITIS EA 4108, site du Madrillet, St Etienne du Rouvray, France
   {yousri.kessentini,thierry.paquet}@univ-rouen.fr

   ** Université Européenne de Bretagne, Université de Bretagne-Sud, CNRS, Lab-STICC, Centre de Recherche Yves Coppens, BP 573, F-56017 Vannes cedex, France
   thomas.burger@univ-ubs.fr
   – téléphone : 00 00 00 00 00
   – télécopie : 00 00 00 00 00
   – e-mail : yousri.kessentini@univ-rouen.fr

7. LOGICIEL UTILISÉ POUR LA PRÉPARATION DE CET ARTICLE :

   LaTeX, avec le fichier de style `article-hermes.cls`,
   version 1.2 du 03/03/2005.

8. FORMULAIRE DE COPYRIGHT :

   Retourner le formulaire de copyright signé par les auteurs, téléchargé sur :
   `http://www.revuesonline.com`

SERVICE ÉDITORIAL – HERMES-LAVOISIER
14 rue de Provigny, F-94236 Cachan cedex
Tél : 01-47-40-67-67
E-mail : revues@lavoisier.fr
Serveur web : http://www.revuesonline.com