

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université M'hamed BOUGARA de BOUMERDES



Faculté des Sciences
Département d'Informatique

MEMOIRE DE MAGISTER

Spécialité : Système informatique et génie des logiciels

Option : Spécification de Logiciel et Traitement de l'Information

Ecole Doctorale

Présenté par :

Zouaoui Hakima

Thème

Clustering par fusion floue de données appliqué à la segmentation d'images IRM

Devant le jury composé de:

Dr. MEZGHICHE Mohamed	Prof.	U.M.B. Boumerdes	<i>Président</i>
Dr. MOUSSAOUI Abdelouahab	M.C.	U.F.A. Sétif	<i>Rapporteur</i>
Dr. BABAHENINI Med Chawki	M.C.	U.M.K. Biskra	<i>Examineur</i>
Dr. DJOUADI Yassine	M.C.	U.M.M.Tizi-Ouzou	<i>Examineur</i>

Année Universitaire : 2007/2008

Résumé

Les données traitées en imagerie médicale sont souvent imprécises et/ou incertaines du fait du mode d'acquisition des images ou de la modélisation des connaissances des médecins. Lorsqu'un expert examine une ou plusieurs images médicales, il prend en compte simultanément ses propres connaissances théoriques ainsi que les informations fournies par les images afin d'effectuer son diagnostic. De même, la fusion de données agrège informations numériques et connaissances théoriques et contextuelles afin de fournir une information synthétique pour l'aide au clinicien.

L'objectif de ce mémoire consiste à développer une architecture de fusion de données basée sur la théorie possibiliste pour la segmentation d'une cible à partir de plusieurs sources d'images. Le processus de fusion est décomposé en trois phases fondamentales.

Nous modélisons tout d'abord les informations dans un cadre théorique commun. Le formalisme retenu consiste à faire la coopération entre l'algorithme FCM (C-moyennes floues) dont la contrainte d'appartenance d'un individu à une classe est gérée d'une manière relative et l'algorithme possibiliste PCM (C-means possibilistes) pour les points aberrants .

Nous agrégeons ensuite ces différentes informations par un opérateur de fusion. Celui-ci doit affirmer les redondances, gérer les complémentarités et prendre en compte les conflits soulignant souvent la présence d'une pathologie.

Nous construisons enfin une information synthétique permettant d'exploiter les résultats de la fusion.

Cette architecture développée est mise en oeuvre pour la segmentation des tumeurs cérébrales à partir des images IRM qui comprennent pour l'instant les séquences de base : T1, T2 et densité de protons (DP).

Mots-clés : Data mining, Imagerie médicale, Fusion de données, Segmentation, Cerveau, C-moyennes floues, C-moyennes possibiliste.

Abstract

Information treated in medical imaging is often inaccurate and uncertain due to image acquisition methods or to the modelling of clinician's knowledge. When analysing medical images to make his diagnosis, an expert takes into account his own knowledge as well as the information provided by each image. Data fusion aggregates in the same way numerical, theoretical and contextual knowledge to provide clinicians synthetic information.

The objective of this work consists of developing architecture of information fusion based on the possibilistic theory in order to segment a target from multiple sources of image. The fusion process is divided into three steps:

We first model the available information, numerical or symbolic, in a common theoretical frame. The formalism selected consists in cooperating between the algorithm FCM (Fuzzy C-Means) whose constraint of membership of an individual to a class is relative and the algorithm possibilist PCM (Possibilistic C-means) for the aberrant points.

Then we aggregate these information with a fusion operator. This operator has to affirm redundancy, manage the complementarities and also take into account conflicts that often underline the presence of pathology.

We finally propose a synthetic piece of information that allows to best represent the available data.

This developed architecture is performed for the segmentation of cerebral tumors from MRI images that presently include these routine sequences: T1, T2 and proton density.

Keywords: Data mining, Medical imaging, Tissue characterization, Data fusion, Clustering, Brain, Fuzzy c-means, possibilist c-means.

ملخص

إن المعطيات المعالجة في الصور الطبية غير دقيقة أو (و) غامضة وهذا راجع للصور المكتسبة أو كيفية معالجة الأطباء للمعرفة، عندما يفحص الخبير الصور الطبية يأخذ بعين الاعتبار هذه المعرفة النظرية و كذلك المعلومات المزودة من هذه الصور بهدف تشخيصها، بالإضافة إلى ذلك، المعطيات المحللة هي عبارة عن جمع المعلومات العددية و المعرفة النظرية و السياقية من اجل تزويد الطبيب المختص بالمعلومات التركيبية.

المذكورة تهدف إلى التركيز على تطوير هندسة المعطيات الملتحمة أساسها النظرية الممكنة، انطلاقا من مجموعة مصادر صورية لأجل تقسيم الهدف. معالجة الالتحام تنقسم إلى ثلاثة مراحل أساسية :
في البداية توضع المعلومات في أطار نظري عام، بحيث الطريقة المتبعة تنسق بين خوارزمية المتوسطات الضبابية التي تحسب درجة الانتماء بطريقة نسبية و خوارزمية المتوسطات الممكنة للأفراد (البيكسال) المبعثرة .

وبعدها نقوم بتجميع مختلف المعلومات بواسطة عامل التجميع ،حيث هذا الأخير يؤكد على التكرار و يدير المتكاملات و يأخذ بعين الاعتبار التنازع الناتج في اغلب الأحيان عن وجود المرض.
و في الأخير نقدم معلومات تركيبية شاملة مستنبطة من خلال استغلالنا لنتائج الالتحام.
هذا التطور الهندسي يوضع من اجل تعيين الورم المخي بواسطة صور الرنين المغناطيسي للدماغ .

كلمات المفتاح: التنقيب عن البيانات، صور الرنين المغناطيسي، التحام المعطيات، التقسيم، الدماغ، بين خوارزمية المتوسطات الضبابية، و خوارزمية لمتوسطات الممكنة

TABLE DES MATIÈRES

Introduction générale.....	1
----------------------------	---

Chapitre 1 : Extraction de connaissances et fouille de données

1.1 Introduction.....	4
1.2 Les étapes d'un processus d'extraction de connaissances à partir des données.....	4
1.2.1 Nettoyage et intégration des données.....	5
1.2.2 Pré-traitement des données.....	6
1.2.3 Fouille de données (Data Mining).....	7
1.2.4 Evaluation et présentation.....	7
1.3 Fouille de données (Data mining).....	9
1.3.1 Historique.....	9
1.3.2 Définition.....	11
1.3.3 Principales tâches de fouille de données.....	11
1.3.3.1 La classification.....	11
1.3.3.2 L'estimation.....	12
1.3.3.3 La prédiction.....	12
1.3.3.4 Les règles d'association.....	12
1.3.3.5 La segmentation.....	12
1.3.4 Les méthodes de data mining.....	13
A. Les méthodes classiques.....	13
B. Les méthodes sophistiquées.....	13
1.3.4.1 Segmentation (Clustering).....	14
1.3.4.2 Règles d'association.....	16
1.3.4.3 Les plus proches voisins.....	20
1.3.4.4 Les arbres de décision.....	22
1.3.4.5 Les réseaux de neurones.....	25
1.4 Conclusion.....	29

Chapitre 2 : Méthodes de classification

2.1 Introduction.....	30
2.2 Méthodes supervisées.....	31
2.2.1 Méthodes bayésiennes.....	31
2.2.2 Champs de Markov.....	32
2.2.3 Algorithme des k plus proches voisins.....	33
2.2.4 Réseaux de Neurones.....	34
2.3 Méthodes non supervisées.....	36
2.3.1 Algorithmes de classification non flous.....	37
2.3.1.1 Algorithmes C-moyennes ("Hard C-Means" ou HCM).....	37
2.3.2 Algorithmes de classification flous.....	38
2.3.2.1 Algorithmes C-moyennes flous ("Fuzzy C-Means" ou FCM).....	38
2.3.2.2 Les variantes des C-moyennes flous.....	40
❖ L'approche de Kamel et Selim.....	41

❖	Méthode FCM semi-supervisée	41
❖	FCM et contraintes de voisinage	42
2.3.2.3	Algorithmes les C-moyennes possibilistes ("Possibilistic C-means" ou PCM)	43
2.4	Conclusion.....	49

Chapitre 3 : Fusion de données

3.1	Introduction	50
3.2	Divers contextes théoriques possibles pour la fusion de données.....	51
3.2.1	La théorie des probabilités.....	51
3.2.2	La théorie des croyances.....	52
3.2.3	La théorie des possibilités.....	54
3.2.3.1	Notion de sous -ensemble flou	54
3.2.3.2	Mesure et distribution de possibilité.....	57
3.2.3.3	Mesure de nécessiter	58
3.3	La Fusion de données	59
3.3.1	Définition.....	59
3.3.2	Intérêt de la fusion de données	60
3.3.3	Le processus du fusion de données.....	61
3.3.3.1	Représentation homogène et recalage des informations pertinentes	61
3.3.3.2	Modélisation des connaissances	61
3.3.3.3	Fusion.....	61
3.3.3.4	Décision par choix d'une stratégie.....	61
3.3.4	Classification des opérateurs de fusion	62
3.3.4.1	Opérateurs à comportement constant et indépendant du contexte.....	63
3.3.4.2	Opérateurs à comportement variable et indépendant du contexte	63
3.3.4.3	Opérateurs dépendants du contexte	63
3.3.4.4	Quelques propriétés	64
3.3.5	Fusion en théorie des probabilités	64
3.3.5.1	Combinaison	64
3.3.5.2	Règle de décision	65
3.3.6	Fusion en théorie des croyances	65
3.3.6.1	Combinaison	66
3.3.6.2	Règle de décision	66
3.3.7	Fusion en théorie des possibilistes.....	67
3.3.7.1	Combinaison	67
3.3.7.2	Règle de décision	68
3.4	Conclusion.....	68

Chapitre 4 : Contribution

4.1	Introduction	69
4.2	Modélisation	70
4.2.1	Choix de l'algorithme	70
4.2.1.1	Quel type de méthode ?	70
4.2.1.2	Classification floue ou non floue ?	71
4.2.1.3	C-moyennes floues ou algorithme possibiliste ?	71
1.	Interprétation des degrés d'appartenance	72
2.	Etude comparative entre FCM et PCM	74
4.2.2	Choix des paramètres de l'algorithme	75
4.2.2.1	Initialisation de l'algorithme	76
4.2.2.2	Détermination du nombre de classes	77
4.2.2.3	Choix du paramètre m	77
4.2.2.4	Choix de la distance	77
4.2.2.5	Détermination des paramètres de pondération η_i	79
4.2.2.6	Choix des vecteurs forme	80
4.2.3	Algorithmes utilisés dans notre approche	81
4.3	Fusion	83
4.3.1	Limitations de la fusion probabiliste	84
4.3.2	Comparaison des théories des possibilités et de l'évidence	84
4.3.3	Vers la théorie des possibilités	86
4.4	Décision	87
4.5	Conclusion	89

Chapitre 5 : Résultats et évaluation

5.1	Introduction	90
5.2	Images utilisées	90
5.2.1	Images réelles	90
5.2.2	Fantômes (images de synthèse)	91
5.2.3	Constructions des images simulées	92
5.3	Evaluation et étude comparative	94
5.3.1	Critères de validation	94
5.3.2	Le protocole d'évaluation	96
5.3.3	Évaluations des résultats	97
5.3.4	Comparaison aux C-moyennes floues et C-moyennes possibiliste	99
5.4	Analyse des résultats	101
5.4.1	Images de synthèse	101
5.4.2	Images réelles	105
5.5	Conclusion	108
	Conclusion générale et perspectives	109
	Annexe - Eléments d'anatomie cérébrale et IRM	112
	Bibliographie	125