

INSTITUT FEMTO-ST

DÉPARTEMENT D'INFORMATIQUE DES SYSTÈMES COMPLEXES (DISC)

Etat de l'art sur la segmentation d'images par intelligence artificielle

EQUIPE DEODIS

Directeur de thèse : Pr JEAN-CHRISTOPHE LAPAYRE
Encadrant : Dr JULIEN HENRIET

Florent MARIE

Table des matières

Introduction	3
1 Localisation automatique de germes dans une image pour une segmentation par croissance de régions	4
1.1 Prince Pal Singh, Jaswinder Singh, "Automatic seed placement in region growing image segmentation", Journal of Engineering, Computer and applied sciences, Vol 2, No 7, July 2013	4
1.2 Shweta Kansal et al, "Automatic Seed Selection Algorithm For Image Segmentation Using Region Growing", International Journal of Advances in Engineering and Technology, June 2015	4
1.3 Mohammed Abdelsamea, "Automatic seeded region growing for 2D biomedical images segmentation" . .	5
1.4 Malek et al, "Seed Point Selection for Seed-Based Region Growing in Segmenting Microcalcifications", September 2012	6
1.5 Fan et al, "An automatic algorithm for sementic objec generation and temporal tracking", Signal Proces- sing Image Communication (2002) 145-164	6
2 Raisonnement à Partir de Cas	7
2.1 Approche générale	7
2.1.1 Perner, "Why Case-Based Reasoning is Attractive for Image Interprétation", Alha and I. Watson (Eds), Case-Based Reasoning Research and Developments, Springer Verlag, 2001	7
2.1.2 Perner, "Image processing in Case Based Reasoning", The Knowledge Engeneering Review, Vol. 20 :3, 311-314, 2006	8
2.1.3 Perner, "Introduction to Case Based Reasoning for Signals and Images", Case Based Reasoning on Signals and Images, Springer, Chap 1, 1-24, 2007	8
2.1.4 Perner, "An architecture for a CBR image segmentation system", in K.-D. Althoff, R. Begrmann and L. K. Brantings (Eds). Case-Based Reasoning Research and Development (pp. 525-535), 1999. Berlin : Springer Verlag	8
2.1.5 Frucci, Perner and Sanniti Di Baja, "Case-Bases-Reasoning for Image Segmentation", Interna- tional Journal of Pattern Recognition and Artificial Intelligence, Vol. 22, No. 5 (2008) 829-842. World Scientific Publishing Company	9
2.2 Calcul de la similarité entre deux images	10
2.2.1 Zamperoni, P., Starovoitov, V., 1995. How dissimilar are two gray-scale images. In : Proceedings of the 17th DAGM Symposium.Springer, Berlin, pp. 448455	10
2.2.2 Perner, "An architecture for a CBR image segmentation system", in K.-D. Althoff, R. Begrmann and L. K. Brantings (Eds). Case-Based Reasoning Research and Development (pp. 525-535), 1999. Berlin : Springer Verlag	10
2.3 Evaluation automatique d'un résultat de segmentation sans verité terrain	10
2.3.1 Zhang, S., 1997. Evaluation and comparision of dierent segmentation algorithm. Pattern Recog- nition Letters 18 (10), 963968	10
2.3.2 Philipp Foliguet, Guingues. "Evaluation de la segmentation d'images : état de l'art, nouveaux indices et comparaisons". 2005	11
2.4 Organisation de la base de cas	12
2.5 Maintenance de la base de cas	12

3 La segmentation d'images avec des réseaux de neurones	13
3.1 Convolutional Neural Network	13
3.1.1 Ciresan D. et al, "Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images". Nips, p 1-9, 2012	13
3.2 Full Convolutional Network	14
Conclusion	15
Bibliographie	17

Introduction

Cet état de l'art a lieu dans le contexte du projet SAIAD (Segmentation Automatique par Intelligence Artificielle Distribuée) qui consiste à créer un système d'intelligence afin de segmenter automatiquement des images médicales sans intervention humaine ou, à minima, avec une intervention très limitée de la part d'un opérateur. La segmentation est un procédé consistant à isoler un ou plusieurs objets d'intérêt sur une image.

Le but médicale derrière ce projet d'envergure est de rendre systématique l'exploitation des informations apportées par l'imagerie médicale dans le cadre du traitement chirurgicale des tumeurs du rein chez l'enfant. Ces images sont en effet un support capital pour orienter au mieux la chirurgie (et même plus généralement la thérapie) et décider si l'organe touché par la tumeur doit être enlevé ou non. Cependant, le temps important que représente le traitement des images les rends difficilement exploitables en routine. La conséquence directe est que ces informations sont le plus souvent ignorés et une chirurgie totale (ablation de l'organe) est pratiqué par défaut ; une perte dommageable pour le patient.

Pour réaliser une segmentation, les procédés d'apprentissages machines semblent être une méthode efficace ouvrant la voie à une automatisation complète du processus. Cet état de l'art porte sur certaines de ces méthodes, en particulier les réseaux de neurones et le raisonnement à partir de cas. D'autres techniques permettant d'automatiser les segmentations sont également abordées.

Chapitre 1

Localisation automatique de germes dans une image pour une segmentation par croissance de régions

La croissance région est une technique de segmentation d'image très utilisée. Son principe consiste à désigner une série de pixels comme des germes. Ces germes forment des régions initiales que l'algorithme va faire grandir par aggrégation des pixels environnants suivant la satisfaction d'un critère. Une fois que l'algorithme a atteint son point de convergence, les différentes régions forment la segmentation finale. L'inconvénient majeur de cette technique se trouve dans son étape d'initialisation qui consiste à choisir les pixels utilisés comme germes. Cette étape est manuelle (dans une croissance régions classique) et pose donc un problème pour automatiser le processus. L'état de l'art de ce chapitre porte sur les moyens existants pour automatiser la sélection des germes.

1.1 Prince Pal Singh, Jaswinder Singh, "Automatique seed placement in region growing image segmentation", Journal of Engineering, Computer and applied sciences, Vol 2, No 7, July 2013

Résumé de l'article

Les auteurs détaillent une méthode de segmentation par croissance de régions. La croissance de région utilise un critère semblant se rapprocher d'un gradient. Si l'intensité d'un pixel candidat est trop éloigné (changement abrupte par rapport à la région), alors il est rejeté. L'initialisation de la croissance région se fait en déterminant une région d'intérêt et en calculant le barycentre de cette région.

Intérêt de l'article pour le projet

Aucun détail n'est donné sur la première étape consistant à déterminer la zone d'intérêt où l'emplacement de la germe est calculé. L'article se contente de dire que cette est déterminée suivant "le fond et les propriétés de l'objet dans l'image en niveau de gris", ce qui le rend peu intéressant. De plus, ceci amène à se demander si cette étape de détermination de la ROI est manuelle (par des méthodes morphologiques utilisateurs dépendantes), ce qui enlèverait tout l'intérêt à la méthode dans le cadre de notre travail.

1.2 Shweta Kansal et al, "Automatic Seed Selection Algorithm For Image Segmentation Using Region Growing", International Journal of Advances in Engineering and Technology, June 2015

Résumé de l'article

L'article reprend une méthode de croissance région développée dans un autre papier pour réaliser la segmentation d'images. La croissance région en elle-même n'a rien de particulier, à ceci près qu'elle élimine les régions ne satisfaisant

pas une certaine taille (en les laissant sans label). Le calcul du critère se fait par rapport aux statistiques de la région en question (moyenne). Initialement, la valeur de la “seed” est prise comme moyenne de la région. A chaque fois qu’un nouveau pixel remplit le critère et est incorporé dans la segmentation, la moyenne de la région est recalculée et sert de critère pour les futures incorporations. Afin d’initialiser correctement et automatiquement cette segmentation, l’article propose un algorithme qui va déterminer automatiquement l’emplacement des “seeds” dans l’image pour la mener à bien. A la fin du placement des “seeds”, il doit y en avoir autant que de régions à segmenter dans l’image. Le placement des “seeds” se base sur leur fréquence d’apparition après un processus de fusion (“merge”).

L’algorithme est le suivant :

Etape 1 : Classement des pixels par ordre croissant de niveau de gris

Etape 2 : Calcul des fréquences pour chaque niveau de gris (histogramme)

Etape 3 : Utiliser le pixel n1 comme premier germe

Etape 4 : Parcourir les voisins du germe et les fusionner pour obtenir une germe améliorée :

- Si le critère de fusion est rempli
 - Remplacer la valeur du germe par la moyenne des pixels fusionnés
 - Mettre à jour la fréquence pour la nouvelle valeur du germe
 - La position du germe n’est pas modifié
- Autrement
 - Le pixel est utilisé comme nouveau germe
- Répéter l’étape 4 tant que tous les pixels ne sont pas fusionnés à une région

Etape 5 : Trier les seeds par ordre décroissants de fréquence d’apparition

Les résultats fournis par les auteurs semblent bons. Le programme est parvenu à segmenter la totalité des objets présents dans l’image avec une précision qui semble tout à fait correcte même pour des objets ayant une forme relativement complexe.

Intérêt de l’article pour le projet

L’intérêt de ce travail se trouve dans l’étape d’initialisation de la croissance région qui permet de placer les “seeds” automatiquement dans l’image sans intervention de l’utilisateur. Ceci représente un grand intérêt pour le projet et permettrait donc de résoudre le problème d’initialisation de la croissance région dans le but de permettre une automatisation par un processus IA. Cependant, il faut noter que les images utilisées pour tester l’algorithme sont des images où la distinction entre les objets à segmenter et le background est très claire. Dans le contexte d’images médicales, où le fond et les organes sont souvent peu distinguables et noyé dans le bruit, les résultats sont soumis à caution.

1.3 Mohammed Abdelsamea, "Automatic seeded region growing for 2D biomedical images segmentation"

Résumé de l’article

L’auteur présente une nouvelle méthode pour initialiser un algorithme de croissance région pour faire de la segmentation d’images médicales (issues d’imagerie cellulaire dans son cas). Le but du papier est de proposer une approche pour permettre de placer automatiquement des seeds dans une image, sans intervention de l’utilisateur.

Pour ce faire, la méthode requiert plusieurs étapes :

- **Débruitage** de l’image à l’aide d’un filtre passe bas (filtre moyenneur par exemple)
- **Seuillage** grâce à la méthode d’Otsu qui calcule automatiquement un seuil optimal de façon à minimiser la variance intraclasse. Cette étape permet d’obtenir une image binaire contenant toutes les seeds potentielles.
- L’auteur utilise ensuite une approche par la **technique des K-moyennes** pour “segmenter” la segmentation en plusieurs zones d’intérêt distinctes. Tous les pixels sélectionnés par le seuillage forment un ensemble de seeds potentielles à utiliser dans la croissance région. Ces pixels sont extraits de l’image et analysés par un algorithme d’apprentissage (K-moyennes) afin de les séparer en plusieurs groupes, ce qui correspondrait à plusieurs zones d’intérêt.
- Pour chaque zone d’intérêt, le **barycentre (“centroid”) de la zone est calculé** et sa position est utilisée comme seed. On obtient alors une seed pour chaque région d’intérêt extraite de l’image.
- **Les seeds extraites** sont ensuite utilisées dans la croissance région pour servir l’initialisation.

L’auteur parvient à de bons résultats sur des images cellulaires avec de la fluorescence. L’algorithme évite le problème de sur-segmentation et de sous-segmentation. Des résultats qui, en plus d’être automatisés, sont d’après lui de meilleure qualité que les autres approches connues.

Intérêt de l'article pour le projet

Cet article est très intéressant pour mettre au point une technique de segmentation automatique par IA. L'approche défendue par l'auteur trouverait tout à fait sa place dans l'arsenal d'outils que l'on pourrait mettre à disposition de cette IA. Cependant, cette approche n'est pas totalement automatique puisqu'elle a recours à une technique d'apprentissage qui est KMeans. Bien que non supervisée, KMeans nécessite une intervention de l'opérateur pour indiquer combien de classes différentes sont à trouver. Il faut donc soit trouver un moyen de déterminer ce nombre automatiquement, soit réfléchir à la possibilité de remplacer la méthode des KMoyennes par une méthode supervisée (SVM par exemple) pour éviter de devoir spécifier le nombre de classes en sortie. La conséquence est que, dans le cas d'un apprentissage supervisé, il y a nécessité de créer une base d'apprentissage pertinente, ce qui représente en soit un gros travail et ce sans garantie de résultats concluants.

1.4 Malek et al, "Seed Point Selection for Seed-Based Region Growing in Segmenting Microcalcifications", September 2012

Résumé de l'article

Les auteurs présentent une méthode pour pouvoir déterminer automatiquement l'emplacement des seeds pour l'initialisation d'une croissance région. Cette méthode se base sur l'extraction des maxima régionaux afin de déterminer ces emplacements. Chaque maxima est identifié comme une zone d'intérêt. Les "moyennes" des maxima sont utilisées en tant que "seeds" (les auteurs ne précisent pas ce qu'ils entendent par "moyenne"; il s'agit probablement du centroïde de chaque maxima). La procédure décrite par Malek est la suivante :

- **Débruitage** avec un filtre passe bas
- **Extraction** des maxima régionaux
 - Utilisation de la fonction $g=f-1$
 - reconstruction géodésique de g par f
 - Différence entre f et g
- **Calcul du centroïde** pour chaque maxima
- **Placement d'une seed** à chaque centroïde calculé

Intérêt de l'article pour le projet

Cet article présente une nouvelle façon de placer des seeds automatiquement dans une image. Cependant, l'article n'est pas toujours claire sur le but de chaque étape et il reste encore des zones d'ombres. Il reste toutefois un papier intéressant qu'il faudra garder en tête.

1.5 Fan et al, "An automatic algorithm for semantic object generation and temporal tracking", Signal Processing Image Communication (2002) 145-164

Résumé de l'article

La papier vise à développer une méthode de détection d'humains dans un flux vidéo. Pour ce faire, il se base sur une segmentation automatique de l'image par détection des contours, puis par un processus de croissance de région. Ces deux étapes permet d'effectuer par la suite une détection de visages.

Intérêt de l'article pour le projet

Il s'agit d'un travail intéressant qui trouvera peut être son intérêt pour le projet SAIAD, qui pourra s'appuyer sur leur méthode pour développer une approche de segmentation pour le modèle multi-agents. Cependant, il n'a pas beaucoup d'intérêt en matière de placement des seeds dans l'images. Le processus de croissance région ne s'appuie pas sur des seeds à placer dans l'images. D'ailleurs, il ne permet pas une véritable segmentation d'un objet, mais permet de récupérer tous les contours dans une image.

Chapitre 2

Raisonnement à Partir de Cas

Le raisonnement à partir de cas est une technique d'IA permettant à un système de résoudre un problème nouveau en faisant appel à son expérience passée. C'est une méthode très puissante dès lors que l'on souhaite concevoir un système automatisé mais souffrant d'un certain nombre d'inconvénients (du moins actuellement) empêchant d'atteindre cet idéal. L'état de l'art suivant porte en particulier sur le RàPC appliqué au traitement d'image (surtout concernant la segmentation), mais certains articles plus généraux ou avec une application intéressante à un autre domaine que l'imagerie sont examinés.

Il existe plusieurs points particuliers dans un système de RàPC qui forment autant de verrous techniques à étudier :

1. Comment calcule-t-on la similarité entre deux images ?
2. Comment évaluer la justesse d'un résultat en absence de vérité terrain ?
3. Comment adapter les paramètres d'une segmentation à partir d'une solution proche ?
4. Comment organiser la base de cas ?
5. Comment gérer la maintenance de la base de cas ?

Les sections de l'état de l'art sur le RàPC seront structurées selon ces verrous techniques. La première section concernera des articles généraux ne pouvant pas rentrer spécifiquement dans aucun de ces domaines.

2.1 Approche générale

2.1.1 Perner, "Why Case-Based Reasoning is Attractive for Image Interpretation", Alha and I. Watson (Eds), Case-Based Reasoning Research and Developments, Springer Verlag, 2001

Résumé de l'article

Il s'agit d'une sorte de passage en revue des travaux sur l'interprétation d'images grâce à un système basé sur le Raisonnement à partir de Cas (RàPC). Le problème des programmes de traitement d'images est leur manque de capacité à pouvoir interpréter des images pour en extraire de l'information abstraite de haut niveau. Le RàPC est une méthode étudiée en vue de donner du sens aux images pour un ordinateur. Cela passe par plusieurs étapes au sein d'un système d'interprétation d'images. Ce système consiste à extraire d'une image des caractéristiques simples ("low features") pour, en les combiner, créer des caractéristiques abstraites ("symbolic features"). Chaque "symbolic features" est décrite comme la combinaison de plusieurs "low features". Une fois les "symbolic features" extraites de l'image, celle-ci peut être comparée à une base de données par un processus de classification pour en déterminer le contenu.

Intérêt de l'article pour le projet

Ce papier est très riche en information. Il ne donne pas de détail très technique sur la façon de mettre en œuvre un traitement d'images par RàPC mais donne le concept d'un système d'interprétation d'images. L'article est dense, et fait appel à un grand nombre de notions. Il permet d'obtenir une vue d'ensemble du sujet. Il donne également un grand nombre de références sur le sujet (interprétation d'images et/ou RàPC).

2.1.2 Perner, "Image processing in Case Based Reasoning", The Knowledge Engineering Review, Vol. 20 :3, 311-314, 2006

Résumé de l'article

Perner nous donne un aperçu d'ensemble de l'utilisation de la RàPC dans le domaine du traitement d'images en citant un grand nombre de travaux sur le sujet, et ce dans plusieurs sous catégories. Ci-dessous sont indiqués les références qui n'ont pas encore été inclus à l'étude bibliographique sur le sujet.

Grimnes and Aamodt (1996) : Système RàPC pour l'interprétation d'images CT-Scan

Micarelli et al (2000) : Application de la RàPC à la reconnaissance d'objets dans une image (avec une approche ondelette)

Perner (2000) : Système RàPC optimisant la segmentation dans des images suivant les conditions d'acquisition et la qualité des images (images US)

Perner (2002) : Système alliant RàPC et classification par dissimilarité pour la reconnaissance de modèles

Perner et Bühring (2004) : Travail pour rendre plus robuste face aux variations du modèle les systèmes de reconnaissance d'objets par RàPC

Holt et Benwell (1996,1999) : Combinaison de la RàPC avec un autre système pour exploiter l'information spatiale

Carswell et al (2002), O'Sullivan et al (2004,2005) : Utilisation de mesures de similarité spatiales pour comparer la localisation spatiale d'objets dans des photos

Intérêt de l'article pour le projet

Cet article présente une vue d'ensemble de la RàPC dans le traitement d'images, ce qui le rend très intéressant et permet de faciliter le travail sur un état de l'art du domaine. Cependant, ce travail de synthèse est ancien (il date de 2006) et ne couvre donc pas les 10 dernières années.

2.1.3 Perner, "Introduction to Case Based Reasoning for Signals and Images", Case Based Reasoning on Signals and Images, Springer, Chap 1, 1-24, 2007

Résumé de l'article

Petra Perner publie ici est un article généraliste sur la RàPC, constituant la première partie du livre (introduction), où sont abordés de manière assez général différents concepts importants : principe, organisation de la base de cas, calcul de la similarité, maintenance de la base de cas, représentation d'une image, etc... La quantité d'informations données rend difficile une synthèse dans un état de l'art.

Intérêt de l'article pour le projet

L'article est un condensé d'information qui seront très utiles pour le projet SAIAD. Il permet d'avoir une vue d'ensemble du sujet. A lire en parallèle des autres parties du livres qui traitent chacun des points soulevés dans l'introduction avec plus de détails.

2.1.4 Perner, "An architecture for a CBR image segmentation system", in K.-D. Althoff, R. Begrmann and L. K. Brantings (Eds). Case-Based Reasoning Research and Development (pp. 525-535), 1999. Berlin : Springer Verlag

Résumé de l'article

Perner détaille dans cet article un système de segmentation d'images basé sur la RàPC. Le système est composé de deux unités : Unité de segmentation par RàPC et Unité de gestion de la base de cas.

Unité de segmentation : Pour une tentative de segmentation donnée, le programme va rechercher dans une base de cas une segmentation déjà réalisé avec des caractéristiques proches. Le cas solution le plus proche dans la base du cas problème est sélectionné et ses paramètres sont utilisés pour segmenter l'image d'intérêt. Le module de segmentation ajuste les paramètres pour obtenir la segmentation la plus pertinente. Pour déterminer le cas solution le plus proche du cas problème, une valeur de similarité est calculé entre les deux. Le cas solution de la base ayant la plus haute valeur de similarité est sélectionné.

Unité de gestion : Une fois la segmentation faite, celle-ci est évaluée. Dans l'article, cette évaluation est manuelle par un utilisateur. Cet utilisateur décide si la segmentation est bonne ou non. Si ce n'est pas le cas, un deuxième module de segmentation va chercher à calculer de meilleurs paramètres. Ce module envoie le résultat dans la base de cas pour la compléter (gain d'expérience). Un module intermédiaire vérifie toutefois que ce nouveau cas ne fait.

Intérêt de l'article pour le projet

Cet article offre une vision complète et concrète d'un système de segmentation d'images utilisant la RàPC. Ces informations seront déterminantes pour la conception d'un système multi-agents ayant pour but de faire de la segmentation. Perner aborde également la question du calcul de la similarité entre deux images en proposant sa propre méthode. Cependant, il y a pas mal de zones d'ombre concernant le fonctionnement de l'algorithme :

- Comment réussir à automatiser la phase d'évaluation de la segmentation produite ?
 - En cas de mauvaise segmentation, comment sont calculés les nouveaux paramètres ?
 - Comment le résultat est-il évalué ?
 - Pourquoi seules les mauvaises segmentations semblent incorporées dans la base de cas et non les bonnes ? Réponse : Si un cas stocké permet de parvenir à une bonne solution du problème posé, alors ce cas est suffisant pour parvenir à la résolution et une incorporation dans la base de ce nouveau cas résolu ne permettrait pas d'augmenter son efficacité. Cela augmenterait juste la taille de cette base en y incorporant un doublon.
- De plus, le processus de segmentation en lui-même n'est pas clairement expliqué.

2.1.5 Frucci, Perner and Sanniti Di Baja, "Case-Bases-Reasoning for Image Segmentation", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 22, No. 5 (2008) 829-842. World Scientific Publishing Company

Résumé de l'article

Ces travaux de Perner en coopération avec deux autres chercheuses font directement suite à ceux de 1999 portant sur la conception d'un système de segmentation d'image basé sur un raisonnement à partir de cas. Le fonctionnement est repris, avec toutefois plusieurs différences :

- La technique de watershed est utilisée pour effectuer la segmentation (et non plus celle de la division de l'histogramme)
- Il n'y a plus d'interventions manuelles de la part de l'utilisateur
- Cependant, les parties portant sur l'évaluation des segmentations et sur la gestion de la base de cas (qui étaient très utilisateur-dépendants dans le papier de 1999) ne sont pas inclus dans le système implémenté

La base de cas est construite en segmentant par watershed chacun des cas avec plusieurs paramètres différents donnant plusieurs résultats de segmentation. Le meilleur résultat est retenu par comparaison avec la réalité du terrain (probablement par évaluation visuelle) et les paramètres correspondant sont stockés en tant que meilleurs paramètres de segmentations. Chaque cas dans la base est constitué d'un ensemble de "features" tels que la moyenne, centroïde,... (Perner reprend les features utilisés pour calculer la similitude dans ses travaux de 1999 qui sont d'ailleurs réutilisés également à cette fin) décrivant l'image et les meilleurs paramètres de segmentation associés à cette image pour le watershed. Cette base de cas est ensuite classifiée par un processus d'apprentissage machine. Les cas sont répartis en plusieurs classes. Au sein d'une même classe, on se retrouve donc avec un ensemble de cas ayant chacun leurs propres paramètres de segmentation. Les paramètres de segmentation moyens de cette ensemble forment les paramètres associés à la classe en question.

La similarité utilise un calcul de distance pour chacune de ces features entre le cas problème et l'ensemble des classes de cas solutions. La classe stockée minimisant une distance moyenne de l'ensemble des distances pour chaque feature est considérée comme la classe la plus proche. Ceci se rapproche donc d'un problème de classification. Les paramètres de segmentation associés à cette classe sont utilisés pour segmenter le cas problème.

En ce qui concerne la technique de watershed, elle possède le défaut majeur d'induire une très forte sur-segmentation de l'objet que l'on cherche à extraire de l'image. Cette sur-segmentation peut être limitée par un processus de fusion de région et d'élimination des seeds non pertinente. Perner se base sur le travail précédent de Frucci ("Oversegmentation reduction by flooding regions and digging watershed lines", 2006) qui utilise le calcul de deux critères de similarité entre deux régions. Ces deux critères étaient ensuite comparés respectivement à deux seuils pour chaque région. Ces seuils sont calculés automatiquement à partir des résultats d'un premier watershed. Si une région ne valide pas au moins un des deux tests, alors elle est considérée comme non pertinente et la seed correspondante est enlevée. Puis un nouveau watershed est réalisé, les régions sont testées, etc... Ainsi de suite jusqu'à ce que toutes les régions soient jugées pertinentes. C'est donc un processus itératif.

Puis, dans le papier de 2008, les auteurs précisent cette méthode en ajoutant deux facteurs poids, a et b , qui permettent de ne pas donner le même poids à chacun des deux critères. Elles partent du principe que pour chacune des images, il ne

faut pas donner le même poids à l'un et l'autre des critères. Ces deux critères ne sont plus calculés et testés séparément. Ils permettent de calculer, avec les facteurs a et b ainsi que les deux seuils, un critère de similarité globale qui est comparé à une valeur seuil globale T .

Les variables a , b et T deviennent donc des paramètres de la segmentation.

Les résultats présentés sont encourageants, à la fois pour la technique utilisée (watershed avec les paramètres a , b et T) et le système CBR). En faisant varier les paramètres manuellement, les auteurs parviennent à obtenir une segmentation correcte. Cependant, cette image reste sursursegmentée. Le système de segmentation par RàPC permet d'obtenir des résultats similaires à la méthode manuelle, ce qui est prometteur, mais accentue légèrement le problème de sursegmentation.

Intérêt de l'article pour le projet

En tant que suite des travaux de Perner en 1999, cet article présente un intérêt évident pour le projet SAIAD. Il permet de poser les bases d'un système de segmentation assisté par du RàPC. Cependant, ce système devra obligatoirement être largement amélioré pour être pertinent au sein du projet SAIAD.

Tout d'abord, ce système est incomplet. Perner et ses collègues n'ont en effet pas élaboré la partie portant sur l'évaluation des segmentations et sur la gestion de la base de cas. On peut s'appuyer sur le travail de 1999 où ces parties étaient présentes, mais il s'agissait de parties effectuées manuellement.

Ensuite, la technique de segmentation en elle-même montre ses limites avec une sursegmentation qui ne permet pas une segmentation optimale. De plus, dans le contexte du projet SAIAD où l'on cherche une application à des images médicales, ce phénomène risque d'être amplifié à cause des images bruitées.

En conclusion, l'article est une source d'inspiration de qualité qui aura probablement une grande importance pour la réalisation de SAIAD. Cependant, il ne permet d'obtenir ni un système fonctionnel ni une technique de segmentation avec des performances suffisantes en l'état.

2.2 Calcul de la similarité entre deux images

2.2.1 Zamperoni, P., Starovoitov, V., 1995. How dissimilar are two gray-scale images. In : Proceedings of the 17th DAGM Symposium. Springer, Berlin, pp. 448-455

Résumé de l'article

Intérêt de l'article pour le projet

2.2.2 Perner, "An architecture for a CBR image segmentation system", in K.-D. Althoff, R. Begrmann and L. K. Brantings (Eds). Case-Based Reasoning Research and Development (pp. 525-535), 1999. Berlin : Springer Verlag

Ce papier a déjà été abordé dans la section 2.1.4, page 8 où Petra Perner explique le RàPC dans une approche générale de la segmentation d'images. Mais elle s'intéresse également à la problématique du calcul de la similarité entre deux images de manière plus spécifique.

Résumé de l'article

Intérêt de l'article pour le projet

2.3 Evaluation automatique d'un résultat de segmentation sans vérité terrain

2.3.1 Zhang, S., 1997. Evaluation and comparison of different segmentation algorithms. Pattern Recognition Letters 18 (10), 963-968

Résumé de l'article

Zhang a effectué une comparaison de divers algorithmes de segmentation d'images en les classant en différents groupes suivant les règles suivantes : L'addition de tous les groupes doit couvrir l'intégralité des techniques de segmentation, toutes les techniques d'un même groupe possèdent un attribut commun, toutes les techniques de groupes différents possèdent un attribut différent.

Dans son article Zhang décrit une méthode pour évaluer la pertinence d'une segmentation et ainsi pouvoir comparer des segmentations obtenues par des techniques différentes. Pour ce faire, il test les techniques sur des images de synthèses dont il explicite la conception. Le critère utilisé prend en compte une référence (l'objet qui doit être segmenté) et l'objet obtenu par le processus de segmentation.

Intérêt de l'article pour le projet

Cet article est une déception par rapport à ce qui était espéré (parvenir à avoir une évaluation automatique d'une segmentation quelconque). L'évaluation des segmentations se fait par un critère utilisant une comparaison entre l'objet voulu et l'objet obtenu (on utilise donc une référence de la même manière que des critères plus classiques comme le DICE). Le processus d'évaluation utilisé par Zhang est bel et bien automatique dans le sens où il utilise des objets synthétiques connus (et donc facilement comparable aux résultats de segmentations). Cependant, cette technique ne permet pas d'évaluer des segmentations d'images quelconques (où l'on aurait pas de connaissance à priori de l'objet, donc impossibilité d'avoir une référence). Cependant, les travaux de Zhang, en évaluant différentes techniques de segmentations dans diverses situations, offre une vue complète du domaine de la segmentation d'images, ce qui pourrait être intéressant lorsque l'on concevra l'outil multi agents et que l'on déterminera quelles techniques de segmentation mettre à sa disposition.

2.3.2 Philipp Foliguet, Guingues. "Evaluation de la segmentation d'images : état de l'art, nouveaux indices et comparaisons". 2005

Résumé de l'article

Les auteurs font un état de l'art sur les méthodes afin d'évaluer le résultat d'une segmentation. Une section est consacrée à quelques méthodes s'appuyant sur une vérité terrain (segmentation de référence), mais une seconde présente des approches variées pour effectuer cette évaluation sans recours à cette vérité terrain. Il s'agit en effet une des problématiques majeures du traitement d'images : comment évaluer la pertinence d'une segmentation sans avoir recours à l'intervention d'un utilisateur effectuant une évaluation visuelle ?

Les différentes techniques sont comparées entre elles afin de lister les forces et faiblesses de chacune d'elles.

Les auteurs proposent en plus une méthode énergétique basée sur le calcul d'une énergie séparée en deux termes telle que :

$$E_{totale}(R) = E_d(R) + kE_c(R)$$

R désigne non pas une région, mais une partition. E_d est une énergie d'attache aux données qui calcul la distance entre le modèle et ces données. Elle favorise les régions homogènes. E_c est quant à elle une énergie de complexité pénalisant les segmentations trop fines et les formes complexes et elle une énergie de complexité pénalisant les segmentations trop fines et les formes complexes et pouvant être calculé en tenant compte du périmètre des régions formant la partition R. Le terme k est un coefficient multiplicateur venant donner plus ou moins de poids au terme de complexité. Favoriser E_d donnera des régions bien homogènes mais ayant tendance à sur-segmenter, pouvant même aller jusqu'à une sur-segmentation absolue (dans ce cas là, la région la plus homogène est celle se limitant à un seul pixel).

On peut désigner aussi k sous le terme de paramètre d'échelle. Plus il est faible, plus la segmentation sera morcellée, plus il est élevé, plus la segmentation se grossière (jusqu'à la limite d'une seule région dans l'image).

Trois expressions de E_d sont proposées et comparées (Q, G et D) avec les autres méthodes. Certains critères privilégient les segmentations grossières, d'autres les segmentations plus fines.

Q, G et D ne discriminent qu'assez peu les différents résultats testés. avec pour exception la segmentation floue (qui est légèrement désavantagée) et Kmeans qui a tendance à avoir une énergie interne beaucoup plus importante. Il y a tout de même des différences entre les trois, ce qui démontre que le choix de cette expression a des conséquences.

La conclusion de ce travail est double : **D'une part**, la méthode énergétique offre une bonne évaluation automatique de la qualité de la segmentation. Elle privilégie celles étant le plus en adéquation avec la perception visuelle. Cependant, tout dépend du facteur d'échelle k choisi. Cette méthode (avec l'expression Q de l'énergie interne) privilégie la segmentation floue à de grandes échelles (à cause du terme de complexité) et les autres méthodes (comme la croissance régions) à une échelle faible sur les résultats de segmentation testés. **D'autre part**, les auteurs insistent sur le fait que la qualité d'une segmentation ne peut pas être absolue et est forcément corrélée à un but recherché. Sachant cela, il est difficile d'évaluer cette qualité. Il faudrait trouver un moyen de contextualiser cette évaluation. Le facteur d'échelle k est une première approche pour le faire. Suivant cette logique, il est également difficile de déterminer une méthode d'évaluation qui serait absolument meilleur que les autres. On ne peut qu'observer le comportement de tel ou tel critère. La méthode énergétique proposée offre cependant l'avantage non négligeable d'être adapté en fonction de ce but par l'intermédiaire du facteur d'échelle.

Intérêt de l'article pour le projet

Ce travail est un état de l'art complet, bien qu'assez ancien (2005) sur les moyens d'évaluer une segmentation. Il offre des pistes intéressantes en comparant les critères existants et en proposant même un nouveau critère énergétique. Cependant, il ne permet pas de désigner une meilleure méthode que les autres, tant "une bonne évaluation" d'une segmentation dépend du contexte et du but recherché. La méthode énergétique est un travail intéressant par la possibilité de modifier le comportement de ce critère en tenant compte du but suivi. Il s'agit par contre d'un paramètre laissé à l'appréciation de l'utilisateur, ce qui ne fait que déplacer le problème.

2.4 Organisation de la base de cas**2.5 Maintenance de la base de cas**

Chapitre 3

La segmentation d'images avec des réseaux de neurones

Ce chapitre est consacré à un état de l'art sur l'utilisation des réseaux de neurones pour effectuer de la segmentation d'images. Dans le contexte du traitement d'images, on parle implicitement de réseaux de convolutions dans la plupart des cas. Les réseaux de convolutions permettent d'attribuer à chaque neurone une partie de l'image via un "champ de réception" afin d'effectuer un traitement local de l'information. Chaque couche de convolution produit une ou plusieurs images convoluées appelées aussi "features map".

On distingue plusieurs types de réseaux, notamment les réseaux de convolution classiques (CNN) et les Full Convolutional Network (FCN).

3.1 Convolutional Neural Network

3.1.1 Ciresan D. et al, "Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images". Nips, p 1-9, 2012

Résumé de l'article

Ce papier décrit l'utilisation d'un réseau de convolution pour réaliser une segmentation multiple d'images de membranes cellulaires obtenues par microscopie électronique.

Pour chaque pixel, il y a deux classes possibles : membrane et non membrane. Le réseau calcul pour chacun d'entre eux la probabilité d'appartenir à la classe "membrane". Un score de 1 désigne un pixel étant une membrane, un score de 0 un pixel non membrane. Un score entre 0 et 1 désigne une certaine probabilité d'être une membrane. Le réseau classe chaque pixel en fonction de son environnement en regardant les pixels voisins au sein d'une fenêtre d'une taille fixée par les auteurs. Quand la taille de la fenêtre augmente, le réseau gagne en profondeur car la taille des noyaux de convolution est gardée volontairement faible.

Architecture du réseau :

Le réseau est constitué de 4 x (couches convolution + maxpooling), puis plusieurs couches "fully connected" pour produire un vecteur caractéristique 1D (les différents réseaux utilisés reposent sur ce principe mais le nombre de couche peut varier suivant la taille de fenêtre retenue). La couche de sortie est également fully connected avec un neurone par classe (donc deux). La sortie est post-traitée avec une couche Soft Max pour garantir que chaque sortie peut être interprétée comme une probabilité d'appartenance.

Quatre réseaux légèrement différents nommés N1, N2, N3 et N4 sont entraînés séparément et utilisés en parallèle. Chaque réseau de neurone possède la même architecture mais couvre une fenêtre de taille différente pour la classification des pixels (donnée par le terme w). La moyenne de leurs sorties respectives forme la sortie du système.

Entraînement du réseau :

Le réseau est entraîné sur un ensemble d'apprentissage de 30 images de 512x512. Pour chaque coupe, les pixels étant jugés comme "membrane" sont utilisés comme des exemples positifs (environ 50 000) et le même nombre de pixels "non membrane" sont échantillonnés de manière aléatoire à chaque itération (sans répétition) parmi les pixels "non membrane". Ceci permet de totaliser au total 3 millions d'exemples d'entraînement dans lesquelles les deux classes sont équitablement représentées.

Post-traitement des sorties :

Les auteurs ont remarqués que les résultats sur la base de test étaient faussés en raison du fait que, contrairement à la base d'apprentissage, les deux classes n'étaient pas équitablement représentées. Pour corriger ce problème, ils ont recours à un post-traitement avec l'application d'une fonction polynomiale sur la sortie des réseaux. Il se trouve qu'une fonction polynomiale de degré 3 permet une bonne approximation avec des coefficients calculés par une approche standards des moindres carrés. Une fois déterminée, la même fonction est utilisée pour calibrer les sorties de tous les réseaux entraînés.

Après calibration, les sorties des réseaux sont lissés à l'aide d'un filtre moyen (rayon de 2 pixels). Un seuillage est enfin effectué pour obtenir une segmentation.

Amélioration des performances :

Les auteurs utilisent deux techniques différentes (testées séparément puis couplées) pour améliorer les performances des réseaux en manipulant les données d'entrées : *foveation* et *nonuniform sampling*. Pour une fenêtre donnée, *foveation* consiste à flouter la périphérie afin de forcer le système à tenir davantage compte de la partie centrale (environnement immédiat du pixel à classifier) que de la partie périphérique (intéressante à regarder mais moins pertinente pour classifier le pixel). *nonuniform sampling* consiste à échantillonner la périphérie de la fenêtre avec une résolution décroissante au fur et à mesure que la distance par rapport au centre augmente. Le centre est donc avec la résolution native, alors que la périphérie est sous échantillonnée au fur et à mesure que l'on s'en éloigne.

Calcul de l'erreur :

Les segmentations sont évaluées par un système en ligne qui calcul trois métriques différentes : *Rand error*, *Warping error* et *Pixel error*. Pour une sortie du système, 9 seuillages différents sont réalisés et l'erreur est calculée sur chacun d'entre eux. Le seuillage donnant l'erreur minimum est retenu comme sortie finale du système. Les auteurs ont utilisés *Pixel error* comme fonction de coût du système.

Résultats expérimentaux :

La configuration utilisée est la suivante : i7 950 3.06 GHz, 24 Go de RAM et quatre GTX 580.

Les auteurs ont utilisés pour leurs résultats quatre réseaux avec des tailles w différents : N1, N2, N3 (avec $w = 65$) et N4 (avec $w = 95$). Tous les réseaux ont tournés sur des données avec *foveation* et *nonuniform sampling*. 3 à 5 itérations ont été utilisés pour l'apprentissage pour tous les réseaux. Chaque itération prend un temps variable pour chacun des réseaux (de 170 min pour N1 à 340 min pour N4), ce qui aboutit à un temps total pour l'apprentissage de plusieurs jours. Pour ce qui est de l'utilisation des réseaux après apprentissage, la classification de 8 millions de pixels prend entre 10 et 30 minutes (ce qui signifie entre 30 seconde et 1 minutes pour segmenter une image 512x512).

Les résultats sont comparés à d'autres approches (11 au total) en se basant sur les trois métriques que le système d'évaluation en ligne utilise. L'approche des auteurs permet de meilleurs résultats en minimisant davantage la totalité des trois métriques pour les papiers utilisés comme comparaison, alors même que le système a été calibré uniquement sur *Pixel error* (la minimisation des deux autres métriques est un effet secondaire).

Intérêt de l'article pour le projet

Ce papier est très intéressant pour plusieurs raisons. Tout d'abord, les résultats obtenus semblent très bons et parviennent à surclasser toutes les autres approches de l'époque concourant pour la même compétition. Ensuite, le système fonctionne totalement de manière automatique, ce qui est le but même du projet SAIAD. Enfin, les auteurs utilisent une approche différentes pour l'entraînement en ne considérant pas chaque coupe comme un exemple/individus de la base, mais en prenant chaque pixel accompagnés par ses voisins dans une fenêtre d'une taille donnée. De plus, l'utilisation de plusieurs réseaux différents permet une distribution aisée des tâches pour la diminution du temps de calcul.

3.2 Full Convolutional Network

Conclusion

Table des figures

Bibliographie

- [1] Philippe BESSE. Classification non supervis Cours risr Mr BESSE, professeur Universit Toulouse, Institut Mathtique de Toulouse.
- [2] Omar EL GANAOUI and Matthieu PERROT. Segmentation par la mode des kmeans. [http ://www.tsi.telecom-paristech.fr/pages/enseignement/ressources/beti/hyste-dyn/node2.html](http://www.tsi.telecom-paristech.fr/pages/enseignement/ressources/beti/hyste-dyn/node2.html).
- [3] Mustapha HADD. *Classification de la population en catries socio-nomiques : modologie et application pratique*. PhD thesis, Institut Nationale de Statistiques et d'Economie Appliqu 1999.
- [4] Pascale JARDIN. Analyse statistiques des donn et apprentissage automatique. Cours Institut Supeur de BioSciences de Paris, deuxi ann
- [5] E. LEBARBIER and T. MARY-HUARD. Classification non supervis Cours d'AgroParisTech.
- [6] Adeline PAIEMENT, Majid MIRMEHDI, Xianghua XIE, and Mark C. K. HAMILTON. Integrated segmentation and interpolation of sparse data. *Image Processing, IEEE Transactions on (Volume :23 , Issue : 1)*, pages 110–115, oct 2013.

Abstract