

Engineering's Thesis

To obtain the diploma of **Engineering's Degree**

Field of Study: **Computer Science**

Specialization: **Artificial Intelligence and Data Science**

Theme

Weakly Supervised Brain Tumor Segmentation Using MRI and medical reports

Presented by
GHANDOUZ Amina
Defended on
29/06/2026

Dr. KLOUCHE Badia
Dr. BOUSMAHA Rabab
Dr. DIF Nassima
Mr. BENARIBA Ahmed
Pr. BENSLIMANE Sidi Mohammed
Dr. KECHAR Mohamed
Dr. ABDELHAK Soumia

Présidente
Examinatrice
Encadreur
Co-encadreur
Co-encadreur
Représentant de l'incubateur
Invitée

Academic Year: 2025/2026

Brain tumor segmentation is a critical task in medical image analysis, essential for accurate diagnosis, treatment planning and patient follow-up. Although deep learning has significantly advanced segmentation performance, most state-of-the-art methods rely heavily on large scale and densely annotated datasets which are costly and time consuming to acquire in clinical settings. This has driven the development of weakly supervised learning approaches that aim to achieve strong performance with limited annotation effort.

This engineering thesis introduces Enhanced MedCLIP-SAMv2, a multimodal framework for weakly supervised brain tumor segmentation. Building upon the original MedCLIP-SAMv2 models, the proposed enhancement integrates magnetic resonance imaging (MRI) scans with corresponding clinical reports to leverage both visual and textual information. While MRI provides rich anatomical and structural details, medical reports offer valuable semantic context regarding tumor type, location, size and characteristics.

The framework employs multimodal alignment strategies to transfer knowledge from textual descriptions to the segmentation task enabling effective tumor localization and boundary delineation under weak supervision. Through extensive experiments on publicly available Figshare Brain Tumor Dataset, the Enhanced MedCLIP-SAMv2 demonstrates superior segmentation performance (dice of $\sim 88\%$) compared to existing weakly supervised and multimodal baselines. The results show notable improvements in Dice score validating the effectiveness of the proposed enhancements in bridging the gap between limited supervision and clinical grade accuracy.

This work contributes to the development of more annotation efficient and practical medical image analysis systems, highlighting the potential of multimodal learning in real world clinical scenarios.

Keywords: Brain Tumor Segmentation, Weakly Supervised Learning, Multimodal Learning, Magnetic Resonance Imaging, Medical Reports, Medical Image Analysis, Deep Learning.

La segmentation des tumeurs cérébrales constitue une tâche essentielle en analyse d’images médicales, indispensable pour un diagnostic précis, la planification thérapeutique et le suivi des patients. Bien que l’apprentissage profond ait considérablement amélioré les performances de segmentation, la plupart des méthodes de pointe reposent sur de vastes ensembles de données annotées de manière dense, dont l’acquisition est coûteuse et chronophage en pratique clinique. Cette contrainte a favorisé le développement d’approches d’apprentissage faiblement supervisé visant à obtenir des performances élevées tout en réduisant l’effort d’annotation.

Ce mémoire d’ingénieur présente Enhanced MedCLIP-SAMv2, un framework multimodal dédié à la segmentation faiblement supervisée des tumeurs cérébrales. S’appuyant sur les modèles originaux MedCLIP-SAMv2, l’approche proposée intègre des images d’imagerie par résonance magnétique (IRM) et les comptes rendus médicaux correspondants afin d’exploiter simultanément les informations visuelles et textuelles. Alors que les images IRM fournissent des informations anatomiques et structurelles détaillées, les rapports médicaux apportent un contexte sémantique précieux concernant le type, la localisation, la taille et les caractéristiques de la tumeur.

Le framework proposé met en œuvre des stratégies d’alignement multimodal permettant de transférer les connaissances issues des descriptions textuelles vers la tâche de segmentation, favorisant ainsi une localisation efficace des tumeurs ainsi qu’une délimitation précise de leurs contours dans un contexte de supervision limitée. Des expérimentations approfondies menées sur le jeu de données public Figshare Brain Tumor Dataset montrent que Enhanced MedCLIP-SAMv2 atteint des performances de segmentation supérieures (coefficient de Dice d’environ 88%) par rapport aux approches multimodales et faiblement supervisées existantes. Les résultats mettent en évidence une amélioration significative du coefficient de Dice, confirmant l’efficacité des améliorations proposées pour réduire l’écart entre une supervision limitée et une précision compatible avec les exigences cliniques.

Ce travail contribue au développement de systèmes d’analyse d’images médicales plus efficaces en matière d’annotation et mieux adaptés aux contraintes de la pratique clinique, tout en soulignant le potentiel de l’apprentissage multimodal dans des scénarios cliniques réels.

Mots-clés: Segmentation des tumeurs cérébrales, Apprentissage faiblement supervisé, Apprentissage multimodal, Imagerie par résonance magnétique (IRM), Comptes rendus médicaux, Analyse d’images médicales, Apprentissage profond.

يُعدّ تقسيم أورام الدماغ في الصور الطبية مهمةً أساسية في مجال تحليل الصور الطبية، لما له من دور محوري في التشخيص الدقيق، تخطيط العلاج ومتابعة المرضى. وعلى الرغم من أن التعلم العميق قد حقق تقدماً ملحوظاً في تحسين أداء نماذج التقسيم فإن معظم الأساليب الحديثة تعتمد بشكل كبير على مجموعات بيانات ضخمة ذات تعليقات توضيحية دقيقة وكثيفة والتي يتطلب إعدادها وقتاً وجهداً وتكلفةً مرتفعة في البيانات السريرية. أدى ذلك إلى ظهور أساليب التعلم ضعيف الإشراف التي تهدف إلى تحقيق أداء مرتفع مع تقليل الحاجة إلى عمليات التعليق التوضيحي المكلفة.

يقدم مشروع التخرج الهندسي هذا إطاراً متعدد الوسائط لتقسيم أورام الدماغ باستخدام التعلم ضعيف الإشراف. ويعتمد هذا الإطار على تحسين النموذج الأصلي من خلال دمج صور التصوير بالرنين المغناطيسي مع التقارير الطبية المقابلة لها للاستفادة من المعلومات البصرية والنصية في آن واحد. فبينما توفر صور الرنين المغناطيسي معلومات تشريحية وبنوية غنية، تقدم التقارير الطبية سياقاً دلاليّاً مهماً يتعلق بنوع الورم، موقعه، حجمه وخصائصه.

يعتمد الإطار المقترح على استراتيجيات محاذاة متعددة الوسائط لنقل المعرفة المستخلصة من الأوصاف النصية إلى مهمة التقسيم، مما يتيح تحديد مواقع الأورام ورسم حدودها بدقة حتى في ظل الإشراف المحدود. أُجريت تجارب موسعة باستخدام مجموعة من البيانات العامة وأظهرت النتائج أن الإطار المحسن حقق أداءً متفوقاً في تقسيم الأورام (88.8%) مقارنةً بالأساليب الحالية المعتمدة على التعلم ضعيف الإشراف والأساليب متعددة الوسائط مما يؤكد فعالية التحسينات المقترحة في تقليص الفجوة بين محدودية الإشراف وتحقيق مستوى من الدقة يلائم المتطلبات السريرية.

يسهم هذا العمل في تطوير أنظمة أكثر كفاءة لتحليل الصور الطبية من حيث تقليل الحاجة إلى التعليقات التوضيحية وأكثر ملاءمة للتطبيقات السريرية العملية كما يبرز الإمكانيات الكبيرة للتعلم متعدد الوسائط في السيناريوهات الطبية الواقعية.

الكلمات المفتاحية: تقسيم أورام الدماغ، التعلم ضعيف الإشراف، التعلم متعدد الوسائط، التصوير بالرنين المغناطيسي، التقارير الطبية، تحليل الصور الطبية، التعلم العميق.