جامعة دمشق
كلية الهندسة الميكانيكية والكهربائية
هندسة الحواسيب والأتمتة

تحسين التقطيع الدلالي باستخدام الشبكات العصبونية العميقة

Enhancing Semantic Segmentation Using Deep Neural Networks

م .محمد أنس مخللاتي

د .رائوف حمدان

الملخص

يعمل هذا البحث على تحقيق التقطيع الدلالي وتجزئته على الصور الجوية باستخدام خوارزميات التعلم العميق حيث تم تقديم منهجيتين للعمل، تعتمد الأولى على نماذج UNETالمستخدمة في التقطيع الدلالي واقتراح تابع كلفة خاص لتحسين نتائج التقطيع، بينما تعتمد المنهجية الثانية على النماذج اللغوية الكبيرة و تطبيق الضبط الجيد Fine Tuningلنموذج SAMومقارنة كلا المنهجيتين اعتماداً على معياري سرعة الاستجابة ودقة النموذج، تم هذا البحث باستخدام لغة البرمجة Python وأطر العمل .Tensorflow,Pytorch,Transformers

القسم النظري

التقطيع الدلالي :هو أحد طرق معالجة الصور الرقمية التي تهدف لتحديد المناطق ذات المعنى في الصورة وتمييزها تبعاً لأصناف محددة.

شبكات UNET: هي نوع خاص من الشبكات العصبونية التي تعتمد بنية encoder-decoderلتحليل الميزات في الصورة واعادة بناء خريطة الميزات و قناع الصورة عألي الدقة.

النماذج اللغوية الكبيرة :نوع متطور من الشبكات العصبونية يعتمد بنية Transformersحيث يتم ضبط أحد النماذج التي تعتبر على أنها Foundation Modelsوتوجيهه لتحقيق المهمة الخاصة بالبحث.

القسم العملي

المنهجية الأولى المقترحة :تعتمد في المرحلة الأولى على المعالجة المسبقة للبيانات من حيث زيادة عدد العينات وتهيئتها ومن ثم في المرحلة الثانية بناء نموذج UNETالمكون من 9طبقات بالمجمل، 4منها للمرمز و 4لمفكك الترميز وطبقة الربط بينهما.

تم في هذه المنهجية تنجيز أربعة اختبارات لمقارنة نتائج تطبيق تابع الكلفة المقترح مع بعض التوابع الأخرى ودراسة نتائج كل اختبار على حده.

	دقة التدريب	خسارة التدريب	دقة التقييم	خسارة التقييم
الاختبار الأول CE with val. Data	0.83	0.484	0.841	0.566
الاختبار الثاني CE	0.88	0.037	0.85	0.0468
الاختبار الثالث MAE	0.89	0.034	0.86	0.045
الاختبار الرابع dice-loss	0.87	0.91	0.86	0.91

القسم العملي

المنهجية الثانية المقترحة :تعتمد على الضبط الجيد fine-tuningلنموذج segment anythingالمخصص لتنجيز عملية التقطيع الدلالي حيث يتم تحويله من نموذج عام لنموذج قادر على تنجيز التقطيع الدلالي للصور الجوية التي تم اعتمادها في هذا البحث.

تم في هذه المنهجية اختبار أداء النموذج اعتماداً على البيانات الخاصة بالمنهجية الأولى وكذلك تابع الكلفة المقترح لمقارنة أداء المنهجيتين.

	العتاد	الدقة	زمن الاستجابة
المنهجية الأولى U-Net Approach	GPU, 15 GB RAM	0.86	22ms
المنهجية الثانية LLMs Approach	GPU, 15 GB RAM	0.94	1170ms

النتائج والمناقشة

تمت دراسة التقطيع الدلالي وتجزئته على الصورة الجوية الملونة وفق منهجيتين:

المنهجية الأولى ببناء نموذج UNETمكون من تسعة طبقات و واعتماد تابع كلفة خاص Soft-Dice Lossلتحسين عملية تجزئة التقطيع

المنهجية الثانية اعتماداً على نموذج SAMو ضبطه الجيد باستخدام تابع الكلفة الخاص نفسه وعدد أقل من عينات التدريب.

النتائج :يظهر اختبار المنهجيتين أن استخدام نموذج UNETبعد معالجة الصور الجوية بطريقة خاصة يملك القدرة على تجزئة التقطيع الدلالي بدقة تصل 86%وبتأخير زمني صغير نسبياً حيث يمكن اعتماد نتائج المنهجية الأولى في تطبيقات الزمن الحقيقي، بينما تظهر المنهجية الثانية أن الضبط الجيد لنموذج لغوي كبير يحقق دقة أعلى في عملية تجزئة التقطيع الدلالي 0.94%ولكن بتأخير زمني كبيرة نسبياً وبالتالي يمكن استخدام المنهجية الثانية في التطبيقات التي لايكون فيها الزمن عاملاً حرجاً .

المراجع

Olaf Ranneberger, Philipp Fischer, and Thomas Brox, 2016-U-NetConvolutional Networks for Biomedical Image Segmentation

.Alexander Kirillov1 et all,2023, Segment AnyThing(SAM), Meta AI Research

Fang chen, Ning Wang, Bo Yu, Lei wang, 2022- Res2-Unet, a new Deep Learning, architecture for building detection from high spatial resolution images