

Semantic Segmentation of Aerial Images Using Fusion of Deep Learning Models on Satellite and Drone Datasets: A Novel Approach

Prepared by

Suhier Nadi Odah

Supervised by

Professor Mohammed Otair

Abstract

New developments in artificial intelligence, including deep learning (DL), have given rise to the new semantic segmentation approach, in which the pixel itself is classified and assigned to a specific class. The main problem with this approach is that it needs a lot of data to learn the models well, and the other problem is the high computational time.

This study proposed an advanced semantic segmentation approach for aerial imagery using a fusion of U-Net and U-Net++ models. The main experiment is based on the fusion of the best architectures by fusing the strengths of the individual models U-Net and U-Net ++ used with several deep learning models as their backbone to improve performance. To achieve that, a new mask-based fusion method is proposed by taking into account all segmentation masks of the best models and applying the intersection between each of the two corresponding masks, then choosing the fused mask with the best mIoU value to get the best segmentation results with the lowest error rate.

For the drone dataset, which contains 23 categories, the best pixel accuracy was (89.25%) resulting from the fusion of the three models (U-Net MobileNetV2, U-Net EfficientNet, and U-Net ResNet50), and the best

coefficient (mIOU) was (0.5143) resulting from the fusion (U-Net and U-Net++).

For the aerial dataset, which contains 6 categories, the best pixel accuracy was (88.08%) resulting from the fusion of the four models (U-Net MobileNetV2, U-Net VGG16, U-Net EfficientNet, and U-Net ResNet50), and the best result of the coefficient (mIOU) was (0.671%) resulting from the fusion of (U-Net and U-Net++), which resulted in significant improvements in the segmentation performance compared to the individual models and previous state-of-the-art approaches.

A detailed comparison table was provided showing the advantages of the proposed model in terms of methodology, dataset size, and performance metrics compared with other relevant studies.

التجزئة الدلالية للصور الجوية باستخدام دمج نماذج التعلم العميق في مجموعات بيانات

الصور الجوية والطائرات بدون طيار: نهج جديد

إعداد

سهير نادي عوده

إشراف

الأستاذ الدكتور محمد عطير

الملخص

أدى التطور الجديد في الذكاء الاصطناعي بما في ذلك التعلم العميق (DL) إلى ظهور نهج التجزئة الدلالية الجديد الذي يتم فيه تصنيف البكسل نفسه وتخصيصه لفئة معينة. المشكلة الرئيسية في هذا النهج هي أنه يحتاج إلى الكثير من البيانات لتعلم النماذج جيداً والمشكلة الأخرى في نماذج التعلم العميق هي الوقت الحسابي العالي.

تقترح هذه الدراسة نهجاً متقدماً للتجزئة الدلالية للصور الجوية باستخدام اندماج نماذج U-Net و U-Net++. تستند التجربة الرئيسية إلى اندماج أفضل البنى من خلال دمج نقاط القوة في النماذج الفردية U-Net و U-Net++ المستخدمة مع العديد من نماذج التعلم العميق باعتبارها العمود الفقري لها من أجل تحسين الأداء. لتحقيق ذلك تم اقتراح طريقة دمج جديدة قائمة على دمج نتائج التجزئة Masks تأخذ هذه الطريقة في الاعتبار جميع أفضة التجزئة لأفضل النماذج وتطبق التقاطع بين كل قناعين متقابلين (نموذجي U-Net مختلفين) ثم تختار القناع المدمج الذي يمتلك قيمة mIoU الأفضل للتأكد من الحصول على أفضل نتائج التجزئة وبأقل معدل خطأ.

بالنسبة لمجموعة بيانات الطائرات بدون طيار والتي تحتوي على 23 فئة ، كانت أفضل دقة بكسل هي (89.25%) ناتجة من اندماج النماذج الثلاثة (U-Net MobileNetV2 و U-Net و Net EfficientNet و U-Net ResNet50) وأفضل معامل (mIOU) كان (0.5143) ناتجًا عن اندماج (U-Net و U-Net ++).

بالنسبة لمجموعة البيانات الجوية والتي تحتوي على 6 فئات، كانت أفضل دقة بكسل هي (88.08%) الناتجة عن اندماج النماذج الأربعة (U-Net MobileNetV2 و U-Net و VGG16 و U-Net EfficientNet و U-Net ResNet50) و أفضل نتيجة لمعامل (mIOU) كانت (0.671%) ناتجة عن اندماج (U-Net و U-Net ++). مما أدى إلى تحسينات كبيرة في أداء التجزئة مقارنة بالنماذج الفردية والطرق المتقدمة السابقة وتفوق نماذج الدمج على النماذج الفردية من حيث دقة الاختبار ومعامل (mIOU).

تم توفير جدول مقارنة تفصيلي يوضح مزايا طريقة الدمج المقترحة من حيث المنهجية وحجم مجموعة البيانات ومقاييس الأداء مقارنة بالدراسات الأخرى ذات الصلة.