جامعة دمشق
كلية الهندسة الميكانيكية والكهربائية
قسم هندسة الحواسيب والأتمتة

تطوير آلية للتّعرف على عدّة أشكال في الصّورة باستخدام التّعلم العميق.

Developing a Mechanism to Recognize Several Objects in an Image Using Deep Learning.

إعداد: رنيــم حافــظ كيــوان

الدكتور المشرف: أ.د. محمد مازن المحايري، الدكتور المشرف المشارك: أ.د. قصي كنفاني

الملخص

تمكنت نماذج التعلم العميق لتحديد الكائنات وتصنيفها في الصور من أداء مهمتها على اختلاف قياسات الكائنات وأحجامها واقتطاع أجزاء منها، إلا أنها بالرغم من دقتها وسرعتها لم تمتلك بعد التصور الحقيقي لنظام الرؤية البشرية.يهدف البحث الحالي إلى تقديم إسهام في عملية تطوير نموذج جديد يسهم بتحسين الدقة في أداء هذه المهمة، باستخدام الشبكة Residual Networkبوصفها شبكةأساسية لاستخلاص السمات من صورة الدخل، يُعتمد منها على ثلاث خرائط سمات مختلفة العمق ليُضاف إليها شبكة مكونة من عدة طبقات التفافية من خلال شبكة سمات هرمية مهمتها العمل على تشكيل عدة مستويات من خرائط السمات وبالتالي القدرة على تحديد الكائنات بأحجامها المختلفة، تُستخدم خوارزمية التوجيه الديناميكي التي تعد الأساس لعمل شبكة الكبسولة عند كل مستوى منها في عملية استخلاص السمات أيضا ليتم لاحقا دمج السمات الناتجة عنها إلى خرائط السمات السابقة للوصول لسمات أكثر قوة وثباتاً، وبالتالي القدرة على تحديد واكتشاف عدة كائنات في الصورة باستخدام شبكتي تنبؤ بالمربعات المحيطية ودرجة الصنف بدلا من أن يتم التنبؤ عن طريق شبكة الكبسولة.

القسم النظري

يتكون نموذج الخوارزمية المقترح من جزأين، الأول مستخلص السمات، والثاني يعمل على اقتراح المربعات المحيطية والتصنيف. استُخدمت الشبكة (ResNet)كشبكة استخلاص سمات تتمثل قوتها في إمكانية زيادة عدد الطبقات الالتفافية عمقا باستخدام الوصلات التي تنقل السمات من الطبقات الضحلة إلى الطبقات الأكثر عمقا [1]، [2]. يُستفاد من خرج عدة طبقات من الشبكة المدربة مسبقا باستخدام تقنية نقل التعلم Transfer learning [3]. بحيث تُدخل إلى عدة طبقات التفافية غايتها تجهيز الخرج ليصبح ملائما لتشكيل طبقة الكبسولات الأولية في شبكة الكبسولة [3] , [4] والاستفادة من خوارزمية التوجيه الديناميكي في استخلاص السمات، بحيث نحصل في النهاية على خمسة مستويات لخرائط السمات. اعتُمدت الطبقة conv3_block4_out بحجم خريطة سمات 28x28 على الخرج وعدد خرائط السمات 512، والطبقة conv4_block6_out بحجم خريطة سمات 14x14، وعدد خرائط السمات 1024، والطبقة conv5_block3_out بحجم خريطة السمات 7x7 على الخرج، وعدد خرائط السمات 2048، ومن ثم تم اعتماد عمليات Up-sampling و down sampling والدمج concatenation للحصول على مستويات خرائط السمات.

بعد تشكيل الكبسولات الأولية وتطبيق خوارزمية التوجيه الديناميكي ينتج شعاع يحتوي على السمات الأكثر قوة للكائنات الموجودة في الصورة، أي تقوية أوزان سمات الكائنات الموجودة في الوقت الذي عملت على خفض أوزان الخلفية. تُدمج خريطة السمات الناتجة مؤخرا مع خريطة السمات المدخلة إلى طبقات الكبسولة للحصول على خريطة السمات النهائية وبالتالي الانتقال للمرحلة الثانية بحيث تُمرر خرائط السمات الناتجة إلى شبكتي التنبؤ بالمربعات المحيطية المرجعية Predicted bounding boxesودرجة الصنف التصنيف.

في نهاية مرحلة التنبؤ: تعطي شبكة التنبؤ بالأصناف لكل صورة شعاع التنبؤ بالأصناف [N,80]، وتعطي شبكة التنبؤ بإزاحات المربعات المحيطية شعاع [N,4]، تُدمج النتيجتان ويُحصل على نتيجة النموذج بالتنبؤ بالمربعات المحيطية ودرجة الصنف لكل مربع. نتيجة الدمج هي الشعاع [N,84] أي:

[: , x1,y1, width, height, c1,c2,c3,….,c79,c80]

حيث إن c80 هو درجة الصنف 80، c79 درجة الصنف 79، للمربع الذي تمثلx1,y1, width, height قيم أبعاده المتنبأ بها. وN هي عدد التنبؤات للصورة.

النتائج والمناقشة

1. درس البحث عملية استخلاص السمات من صور قاعدة البيانات COCO باستخدام مفهوم نقل التعلم لنموذج الشبكة ResNet50، وباستخدام خوارزمية التوجيه الديناميكي التي تعد حجر الأساس الذي تقوم عليه شبكة الكبسولة، لزيادة قوة عملية استخلاص السمات.
2. ساهم نموذج ResNet50 من خلال طبقاته الالتفافية الحصول على سمات مكانية أكثر دقة وثبات دون الوقوع في مشكلة تدهور البيانات التي تعاني منها أغلب الشبكات العصبونية الالتفافية عند زيادة العمق.
3. تم استخدام تقنية استخلاص السمات الهرمية في استخدام خرائط سمات محددة من ResNet 50 للحصول على عدة مستويات.
4. تمكن النموذج المقترح من استخدام شبكة الكبسولة في التعرف على عدة أشكال في الصورة وتصنيفها في حين صممت شبكة الكبسولة للتعرف على صنف واحد من كل صورة (استخدمت قواعد بيانات للتعرف على الأرقام، أو الأمراض السرطانية،...).
5. تمكن النموذج من دمج خريطة السمات الناتجة عن السمات الهرمية مع السمات الناتجة عن خوارزمية التوجيه الديناميكي لتقوية السمات والحصول على سمات دلالية مرتبطة السياق، ومن ثم تمريرها إلى شبكتي التنبؤ بالمربعات المحيطية والتنبؤ بدرجة الصنف بدلا من من جعل التنبؤ يتم عن طريق شبكة الكبسولة وهو الأمر المتبع في مثل هذه الأبحاث.
6. قدمت خوارزمية التوجيه الديناميكي أثرا ملحوظا في رفع أداء مهمة التحديد والتصنيف مقارنة مع عدة نماذج تم التوصل إليها في هذا المجال 83.27%.
7. ناقش البحث أداء النموذج عند عدة قيم لعتبة درجة الثقة confidence threshold وحيث أن العتبة 0.5 حققت توازنا في القدرة على تعتيب التنبؤات الصحيحة دون الوقوع في أخطأ FP (التنبؤ بكائنات غير موجودة في الصورة، أو التنبؤ بكائنات مثل الأصناف الموجودة في الصورة ولكن بخلاف التنبؤات الصحيحة) بمتوسط دقة 83.27%.
8. ناقش البحث أداء النموذج عند عدة قيم مختلفة لعتبة التداخل بين المربعات المحيطية حيث حققت العتبة 0.4 متوسط دقة في عملية التنبؤ 60.27%من ناحية انتقاء المربعات المحيطية الأفضل إحاطة.
9. تراوح خطأ التدريب train lossبين 4.016% و3.987%، بينما تراوح خطأ التقييم val_loss بين 3.961 و3.984%.
10. تم بناء مصفوفة الارتباك confusion matrix عند عتبتي confidence threshold 0.4 , 0.5 ودُرست زيادة قيمة الدقة precision عند زيادة قيمة العتبة، مقابل زيادة الحساسية Recall عن تناقص قيمة العتبة.

المراجع

[1]Zhang, A., Lipton, Z. C., Li, M., & Smola A. J. (2021). **Dive into Deep Learning.**, Release 0.17.0, p 287- 293.

[2] He, K., Zhang, X., Ren, S., & Sun J. (2016). **Deep Residual Learning for Image Recognition**. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770–778), 2016.

[3] Zhang, P., Wei, P., & SHuhuan Han, SH. (2020). **CapsNets algorithm**. J. Phys.: Conf. Ser. 1544 012030, doi:10.1088/1742-6596/1544/1/012030.

[4] Sabour, S., Frosst, N., & Hinton, G. E. (2017). **Dynamic routing between capsules**. in Advances in neural information processing systems, 2017, pp. 3856–3866.