

استعمال الشبكات العصبونية الملتفة ونقل المعرفة في التعرف على الصور

إعداد: مؤيد السعدي

إشراف: أ.د. عرفات عوجان

الملخص

يعتبر التعرف على الصور من موضوعات الرؤية الحاسوبية التي تهدف إلى تحديد واحد أو أكثر من الكائنات أو أصناف الكائنات أو المزايا أو النشاطات الموجودة في الصورة المدخلة أو إطارات الفيديو حتى وإن كان محجوبا جزئيا عن المشهد. وتعتبر تقنية الشبكات العصبونية العميقة الملتفة ذروة التقدم فيها هذا المجال لكنها تحتاج للكثير من وقت التدريب وقدرة حسابية عالية.

المقاربة التي تسلكها هذه الرسالة تقترض وجود قيود تتمثل في موارد محدودة من حيث زمن التدريب والقدرة الحسابية العادية المتاحة في الأسواق. إلى جانب ذلك أضيف قيد آخر وهو أن يكون النموذج المطروح قابل للتوسعة. تقنية "نقل المعرفة" استخدمت لإعادة استعمال نماذج مدربة مسبقا متاحة للعموم. إلى جانب تقنيات أخرى لجعلها مجدية وتسريع العملية وجعلها قابلة للتوسعة.

هناك العديد من التطبيقات للتعرف على الصور في نطاقات متعددة مثل التعرف على الوجوه والتعرف على المحارف والتفحص الآلي في خطوط الإنتاج وضبط الجودة والتشخيص الطبي والمهام المتعلقة بالمركبات ذاتية القيادة مثل التعرف على المشاة. إلا أن التطبيق المطروح في هذه الرسالة هو التعرف على الشركة الصانعة للسيارة ونوعها (الموديل) وسنة الصنع من خلال صوتها التي يرفعها البائع على منصة للتجارة الإلكترونية. وتأتي أهمية هذا التطبيق مع تزايد شعبية الهواتف الذكية المزودة بكاميرات حيث أن أخذ صورة أسهل من نقر وصف السيارة حرفا حرفا.

جرى نقل الأوزان من نموذج ImageNet إلى نموذج أول جرى إعادة توليفه ليقوم بتنظيف بيانات التدريب ثم تم عمل عملية نقل أخرى إلى نموذج ثان يستعمل للتعرف على أنواع السيارات المهمة تبعا لحصتها من السوق وبعدها يتم توسعة النموذج كي يشمل نماذج أكثر فأكثر من خلال إجراء تم تقديمه في هذا البحث.

تم إنجاز دقة تزيد عن ٨١٪ في التعرف على ٢٢٩ نوع مختلف من السيارات من خلال إعادة استعمال نموذج Inception V1 المدرب مسبقا لحل مسألة ImageNet ذات الألف فئة بدقة ٦٩,٨٪. ولضمان أن الطريقة المطروحة عامة ولا تتعلق بنطاق معين تم تقييمها من خلال مهام أكاديمية معروفة.