قسم علوم الحاسب
كلية الحاسبات والذكاء الاصطناعى
جامعة بنى سويف

مجلة قسم علوم الحاسب للعلوم المتقدمة

عصر محولات الرؤية (Vision Transformers)

في السنوات الأخيرة، شهد مجال رؤية الكمبيوتر مساهمة كبيرة مع محولات الرؤية (VIT) في عام 2020. ظهرت محولات الرؤية كبنية قوية لمهام التعرف على الصور، مما يشكل تحديًا للشبكات العصبية التلافيفية التقليدية.(CNNs) فى التعليم العميق, نسترسل خلال هذا المقال حول محولات الرؤية، والتعمق في أصولها والتطورات الرئيسية وتأثيرها على مجال رؤية الكمبيوتر.

- **التحولات من البرمجة اللغوية العصبية إلى رؤية الكمبيوتر**

أحدثت بنية Transformer ، التي تم تصميمها مؤخرًا لمهام معالجة اللغة الطبيعية (NLP) في عام 2017، ثورة في هذا المجال بآليات الاهتمام الذاتي والقدرة على التقاط التبعيات طويلة المدى والأرتباط بين الكلمات ومعانيها بغرض افهام الكمبيوتر معنى الجملة، كما ظهر في نماذج الاندماج الحديثة و .GPTولحسن الحظ، تم إثراء مجال الرؤية الحاسوبية بواسطة المحولات من خلال محولات الرؤية التي توسع مفهوم المحولات ليشمل معالجة بيانات الصورة من خلال التعامل مع الصور كمجموعة من الأجزاء الصغيره وفهم ارتباط كل جزء كل من الصوره بالأخر وبالتالى فهم محتوى الصوره. وقد أظهر هذا النهج مؤخرًا قدرته القوية على تحقيق نتائج مماثلة للشبكات العصبية التلافيفية (CNN) فيما يتعلق بمشاكل الرؤية، مثل تصنيف الصور.

- **مما تتكون محولات الرؤية(VIT)**

يتكون ViT من مكونين رئيسيين: جهاز التشفير وجهاز فك التشفير. يتكون برنامج التشفير من مجموعة من طبقات المحولات التي تعالج أجزاء الصور المدخلة، مما يمكّن النموذج من تعلم التبعيات بين أجزاء الصورة على المستوى الجزئى و الكلى. وحدة فك التشفير، التي تكون عادةً طبقة خطية يتبعها دالة SoftMax ، تقوم بإخراج احتمالات الفئة المتوقعة بناءاً على المطلوب.

- **التطورات**

منذ طرحها، أظهرت محولات الرؤية أداءً رائعًا في مهمة تصنيف الصور. ومن الجدير بالذكر أنه تم اقتراح نماذج هجينة تدمج شبكات CNN ومحولات الرؤية، مما يساعد في حل مشاكل الرؤية المعقدة. بالإضافة إلى ذلك، تم بذل الجهود لتكييف محولات الرؤية لمهام رؤية الكمبيوتر الأخرى، مثل اكتشاف الأشياء وتجزئة الصور وفهم الفيديو.

- **التحديات والتوجهات المستقبلية**

أحد القيود هو التعقيد الحسابي، بسبب آلية الاهتمام الذاتي؛ وتشمل الحلول البديلة تقنيات الاهتمام المتناثر وتقريب الانتباه. بالإضافة إلى ذلك، يظل فهم كيفية اتخاذ محولات الرؤية للقرارات وتفسير المعلومات المرئية مجال مفتوح للبحث.

- **خاتمة**

لقد ساهم Vision Transformers في عصر جديد في رؤية الكمبيوتر، مما يشكل تحديًا لهيمنة بنيات CNN التقليدية. ومن خلال تطبيق آلية الاهتمام الذاتي الخاصة بالمحول على الصور، أظهرت Vision Transformers أداءً استثنائيًا في مهام التعرف على الصور.

اعداد

م.م / محمد عرفة

مدرس مساعد علوم الحاسب -كلية الحاسبات والذكاء الاصطناعى – جامعة بنى سويف

رئيس الاصدارات
أ/هشام محمد

رئيس التحرير
ا.م.د/ أحمد النجار

رئيس مجلس الادارة
ا.د /محمد قايد

Address: New Beni-Suef City. Beni-Suef. 62111 Web Site: WWW.fci.bsu.edu.eg
Email: fci@fci.bsu.edu.eg Telephone/Fax: 082 2246796