

| | | | | | | | |
|--------------|-------------|---------|-----------------|---------------|--------------|----------------|------------|
| مقالات علمية | أبحاث علمية | مؤتمرات | فعاليات ونشاطات | براءات اختراع | سير أكاديمية | كورسات تدريبية | تواصل معنا |
|--------------|-------------|---------|-----------------|---------------|--------------|----------------|------------|

| | |
|-----------------------------|---|
| مسابقة أفضل مقالة علمية | تقنيات تحويل النص إلى صورة باعتماد شبكات الخصومة التوليدية (د. زيد قريطم) |
| مقالات كلية الهندسة (حاسوب) | |

ملخص

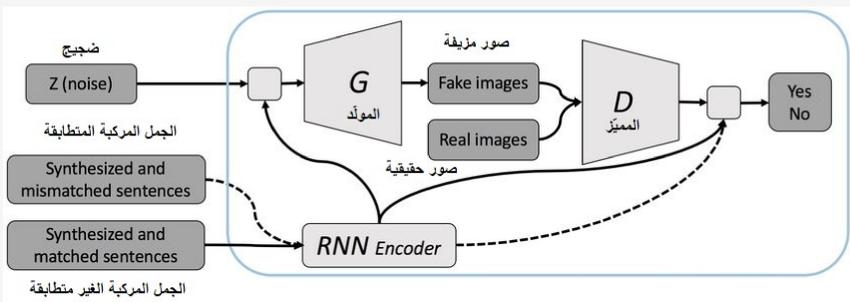
تستكشف هذه المقالة إمكانيات لتطوير شبكات الخصومة التوليدية (Generative Adversarial Networks GANs) في تحويل النصوص إلى صور تعبيرية دقيقة. فبالرغم من التقدم الملحوظ في مجال الذكاء الاصطناعي، لا يزال تحويل النص إلى صورة يمثل تحديًا بسبب التعقيد الكبير في فهم السياق والمعاني المجازية والمشاعر المرتبطة بالنصوص. سنستعرض أهم المنهجيات التي تقوم على دمج الشبكات العصبية العميقة للمعالجة اللغوية مع GANs لتوليد صور فوتوغرافية واقعية تعكس محتوى ومشاعر النص المدخل. كذلك سنقدّم أبرز المقترحات لتطوير وتحسين أداء تلك الشبكات.

الكلمات المفتاحية: شبكات الخصومة التوليدية، تحويل النص إلى صورة، معالجة اللغات الطبيعية، الذكاء الاصطناعي، التعلم العميق.

لمحة عامة

تعد شبكات الخصومة التوليدية (GANs) من أهم أنواع النماذج الحاسوبية في مجال التعلم العميق، والتي تم تقديمها لأول مرة من قبل الباحث إيان جودفيلو وزملائه في عام 2014 [1]. هذا النظام يضم شبكتين عصبونيتين تعملان في تنافس مع بعضهما البعض، وهما: الشبكة التوليدية (Generative Network): هذه الشبكة تحاول إنشاء بيانات جديدة تشبه البيانات الحقيقية. هدفها هو إنتاج بيانات مزيفة لا يمكن تمييزها عن البيانات الحقيقية. الشبكة التمييزية (Discriminative Network): هذه الشبكة تحاول التمييز بين البيانات الحقيقية والبيانات المزيفة التي تنتجها الشبكة التوليدية [1]. هدفها هو الكشف عن البيانات المزيفة بنجاح. تتدرب الشبكتان معاً بشكل تنافسي، حيث تحاول كل شبكة تحسين أدائها استجابة للأخرى. مع تقدم التدريب، تصبح الشبكة التوليدية أكثر مهارة في إنتاج بيانات تشبه الحقيقية، بينما تصبح الشبكة التمييزية أكثر مهارة في التمييز بين البيانات الحقيقية والمزيفة [2]. وتعني بمصطلح الخصومة في تسمية هذا النوع من الشبكات للدلالة على أن هناك نوع من الصراع أو المنافسة بين الشبكتين، وهو ما يدفع كل منهما لتحسين أدائها بمرور الوقت حتى يصبح من الصعب على الشبكة التمييزية التفريق بين البيانات الحقيقية وتلك التي تم توليدها بواسطة الشبكة التوليدية [2]. كما يظهر الشكل (1). تؤدي الطبيعة التنافسية لـ GANs إلى نتائج مثيرة للإعجاب، لكن تدريبها يمكن أن يكون صعبًا ويحتاج إلى الكثير من الموارد الحاسوبية

الشكل (1): بنية شبكات الخصومة التوليدية



خطوات تحويل النص إلى صورة

لتحويل النص إلى صورة باستخدام شبكات الخصومة التوليدية، تتبع الخطوات التالية:

الخطوة 1: معالجة النص: يتم تحليل النص المدخل باستخدام نماذج معالجة اللغة الطبيعية لاستخلاص الكلمات الرئيسية، العبارات المهمة، والمشاعر. تحويل المعاني المستخرجة إلى تمثيلات قابلة للمعالجة بواسطة الشبكة التوليدية.

الخطوة 2: التوليد الأولي للصورة: يأخذ المولد المعلومات اللغوية وينشئ صورة أولية تعكس النص. يتم استخدام بيانات خارجية لتوجيه عملية التوليد وتحسين الدقة.

الخطوة 3: التقييم بواسطة الشبكة التمييزية: تقوم الشبكة التمييزية بتقييم الصور التي تم إنشاؤها من حيث الجودة والدقة السياقية والعاطفية. تعطي ملاحظات للمولد لتحسين الصور في دورات تدريبية متكررة.

الخطوة 4: التكرار والتحسين: يستمر التدريب في دورات حتى تصل الشبكة التوليدية إلى مستوى من الدقة يصعب معه على الشبكة التمييزية التفريق بين الصور المولدة المزيفة والصور الحقيقية.

الخطوة 5: مرحلة ما بعد المعالجة: بعد الحصول على الصورة المقبولة من الشبكة التمييزية، يمكن تطبيق تقنيات ما بعد المعالجة لتحسين الجودة الجمالية للصورة.

دور المولد والمميز: يتلقى المولد تمثيلات النصوص ويحولها إلى صور. يتعلم من التغذية اللفية التي يقدمها المميز لتحسين قدرته على إنشاء صور واقعية. يعمل المميز كحكم لتقييم الصور التي ينتجها المولد. يتم تدريب المولد مما يجعل الصورة تبدو واقعية ومتناسقة مع النص الأصلي. التحدي الرئيسي في هذا النوع من النظم هو تحقيق توليد صورة تعكس بدقة المعاني المعقدة والسياق العاطفي للنص، مما يتطلب دمج معرفة معمقة في المعالجة اللغوية والتوليد البصري [1][2].

أهم الدراسات التي تعتمد GANs

شبكات StackGAN:

- الإنشاء على مرحلتين: قدمت StackGAN عملية جديدة مكونة من مرحلتين. في المرحلة الأولى، يتم إنشاء رسم تخطيطي أساسي للصورة بأشكال وألوان تقريبية من وصف النص. في المرحلة الثانية، تقوم شبكة GAN أخرى بتحسين هذا المخطط وتحويله إلى صورة تفصيلية. يساعد هذا في تقسيم المشكلة المعقدة المتمثلة في إنشاء صور عالية الدقة من أوصاف النص إلى مشكلات فرعية أكثر قابلية للإدارة.
- التعزيز المشروط Conditioning Augmentation: تضيف هذه التقنية تنوعًا إلى تضمين النص المشروط عن طريق إدخال اختلافات صغيرة، مما يجعل النموذج أكثر قوة ويحسن تنوع الصور التي تم إنشاؤها [3].

شبكات AttnGAN:

- آلية الانتباه Attention Mechanism: أضاف AttnGAN آلية انتباه تسمح للنموذج بالتركيز على أجزاء مختلفة من النص عند إنشاء أجزاء مختلفة من الصورة. وهذا مفيد بشكل خاص للأوصاف المعقدة.
- تحسين متعدد المراحل Multi-Stage Refinement: على غرار StackGAN، يقوم AttnGAN بتحسين الصورة على مراحل متعددة. ومع ذلك، فإن آلية الانتباه في كل مرحلة تسمح بمزيد من التحكم الدقيق في عملية إنشاء الصورة، وربط أجزاء من النص بشكل مباشر بتفاصيل محددة في الصورة [4].

شبكات DALL-E:

- بنية المحول Transformer Architecture: على عكس شبكات GAN، يستخدم DALL-E بنية محولات مشابهة لـ

GPT-3. هذا النموذج قادر على فهم وإنشاء كل من النص والصور، مما يسمح له بإنشاء صور من الأوصاف النصية

عن طريق التنبؤ بالبكسلات مباشرة، هذه التقنية تم طرحها من قبل شركة OpenAI.

- التعلم بلمقطة صفرية Zero-Shot Learning: يمكن لـ DALL-E إنشاء صور معقولة من محذلات نصية لم يسبق رؤيتها من قبل (لقطة صفرية)، مما يُظهر مستوى عالٍ من الإبداع وفهم المحتوى [5].

مقترحات لتطوير شبكات الخصومة التوليدية

يمكن تطوير شبكات (GANs) لتحويل النص إلى صور بحيث تشمل:

1- تحسين الفهم السياقي: استخدام نماذج معالجة اللغات الطبيعية المتقدمة مثل GPT أو BERT لتحليل النص واستخلاص المعاني العميقة والمشاعر المرتبطة به.

2- دمج المعرفة الخارجية: إدخال قواعد بيانات معرفية وسيماتيكية لمساعدة الشبكة التوليدية على تحديد العناصر الأكثر صلة لتوليد الصورة.

3- تعزيز التمييز العاطفي: تحسين الشبكة التمييزية لتكون قادرة على تقييم الصور ليس فقط من حيث الواقعية بل أيضاً من حيث التعبير العاطفي والدقة السياقية.

4- تحسين جودة الصورة: تطبيق تقنيات ما بعد المعالجة لتحسين الجودة البصرية للصور التي تم توليدها، مثل الوضوح، الإضاءة، والتفاصيل.

المراجع :

- 1- I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems 27*, Montreal, Quebec, Canada, 2014, pp. 2672/2680.
- 2- I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. New York, USA: MIT Press, 2016.
- 3- Zhang H, Xu T, Li Het al (2017) StackGAN: Text to photorealistic image synthesis with stacked generative adversarial networks. In Proceedings of the Advanced International Conference on Computer Vision (ICCV).
- 4- T. Xu, P. Zhang, Q. Huang, H. Zhang, Z. Gan, X. Huang, and X. He, "Attngan: Fine-grained text to image generation with attentional generative adversarial networks," in CVPR, 2018.
- 5- Ramesh, Aditya; Pavlov, Mikhail; Goh, Gabriel; Gray, Scott; Voss, Chelsea; Radford, Alec; Chen, Mark; Sutskever, Ilya (24 February 2021). "Zero-Shot Text-to-Image Generation". arXiv:2102.12092.

الجامعة الوطنية الخاصة

تأسست عام 2007 و تضم ست كليات :

- كلية طب أسنان
- كلية الصيدلة
- كلية الهندسة (المعلوماتية و الاتصالات)
- كلية الهندسة المدنية
- كلية هندسة العمارة و التخطيط العمراني
- كلية العلوم الإدارية و المالية

مواقع مرتبطة:

🌐 موقع الجامعة الوطنية الخاصة

🌐 موقع المكتبة الرقمية للجامعة الوطنية الخاصة

🌐 موقع الواعة الأكاديمية للجامعة الوطنية الخاصة

🌐 موقع الواحة الطلابية للجامعة الوطنية الخاصة

🌐 موقع بوابة الطلاب الإلكترونية

للتواصل :

📍 سوريا - محافظة حماة - الطريق الدولي حمص حماة

☎️ 0096334589094

☎️ 00963335033

✉️ info@wpu.edu.sy