



بحث بعنوان:
"رؤية الكمبيوتر والتعلم العميق"
مقدم كنشاط لمقرر الذكاء الاصطناعي

اعداد الطالب
الوليد فهد عبده محمد الجماعي

اشراف المهندس:
ياسر الشاعري

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

المقدمة:

تعد رؤية الكمبيوتر والتعلم العميق من أهم التقنيات الحديثة التي غيرت بشكل جذري الطريقة التي تتفاعل بها الآلات مع العالم المحيط بها، إذ تمكن رؤية الكمبيوتر الأنظمة من تفسير وفهم الصور والفيديوهات بطريقة مشابهة للبشر، بينما يوفر التعلم العميق الأساس النظري والتقني لتحقيق هذا الفهم العميق والمعقد. هذا البحث يهدف إلى استعراض الأسس النظرية لرؤية الكمبيوتر والتعلم العميق، واستكشاف العلاقة بينهما، والتطبيقات العملية، بالإضافة إلى مناقشة التحديات والآفاق المستقبلية في هذا المجال.

رؤية الكمبيوتر: تعريف وأهمية

تعريف رؤية الكمبيوتر

رؤية الكمبيوتر هي فرع من فروع علوم الحاسوب يركز على تمكين الآلات من "رؤية" العالم المحيط بها وتحليل الصور والفيديوهات لفهم محتواها. الهدف الأساسي هو تطوير أنظمة تستطيع تفسير البيانات البصرية بطريقة تمكنها من اتخاذ قرارات أو أداء مهام محددة بشكل مستقل.

أهمية رؤية الكمبيوتر

تكتسب رؤية الكمبيوتر أهمية كبيرة في العديد من المجالات مثل:

- **الصناعة والتصنيع:** استخدام الروبوتات القادرة على التعرف على الأجسام والقيام بعمليات التجميع والتفتيش.
- **الرعاية الصحية:** تحليل الصور الطبية مثل الأشعة السينية والتصوير بالرنين المغناطيسي للكشف عن الأمراض.
- **الأمن والمراقبة:** تحليل الفيديوهات في أنظمة المراقبة للكشف عن السلوكيات المشبوهة أو التعرف على الوجوه.
- **السيارات ذاتية القيادة:** استخدام الرؤية الكمبيوترية لفهم البيئة المحيطة واتخاذ قرارات القيادة الآمنة.

التعلم العميق: نظرة عامة

تعريف التعلم العميق

التعلم العميق هو فرع من فروع التعلم الآلي يعتمد على الشبكات العصبية الاصطناعية ذات الطبقات المتعددة (الشبكات العصبية العميقة) لمعالجة البيانات واستخراج الميزات المعقدة منها. يتميز التعلم العميق بقدرته على التعلم من كميات ضخمة من البيانات وتحقيق أداء عالٍ في مهام مثل تصنيف الصور، التعرف على الصوت، والترجمة الآلية.

مكونات التعلم العميق

تشمل مكونات التعلم العميق:

- **الشبكات العصبية الاصطناعية**: تتكون من طبقات من العقد العصبية التي تتصل ببعضها البعض، حيث تقوم كل طبقة بمعالجة جزء من المعلومات وتقديمها للطبقة التالية.
- **التعلم بالإشراف وغير الإشراف**: يمكن تدريب الشبكات العصبية باستخدام بيانات مصنفة (إشرافية) أو غير مصنفة (غير إشرافية).
- **التحسين والتدريب**: تشمل تقنيات مثل الانحدار التدريجي العكسي لتحسين أوزان الشبكة وتقليل خطأ التنبؤ.

العلاقة بين رؤية الكمبيوتر والتعلم العميق

التكامل بين التقنيتين

التعلم العميق يعتبر أداة قوية في مجال رؤية الكمبيوتر، حيث يمكن للشبكات العصبية العميقة اكتشاف الميزات المعقدة والتمثيلات البسيطة من البيانات البصرية دون الحاجة إلى استخراج يدوي للميزات. هذا التكامل بين الرؤية الكمبيوترية والتعلم العميق أدى إلى تحسين كبير في دقة وكفاءة الأنظمة البصرية.

أمثلة على استخدام التعلم العميق في رؤية الكمبيوتر

- **الشبكات العصبية الالتفافية (CNNs)**: تستخدم بشكل واسع في تصنيف الصور، حيث تستطيع استخراج الميزات المكانية من الصور بفعالية.
- **شبكات التحويل (Transformers)**: بدأت تستخدم في رؤية الكمبيوتر لتحسين أداء المهام مثل التعرف على الأشياء وترجمة النصوص في الصور.
- **التعلم الانتقالي**: يسمح بنقل المعرفة المكتسبة من مهام معينة إلى مهام أخرى، مما يقلل من الحاجة إلى كميات ضخمة من البيانات الجديدة.

النماذج والتقنيات المستخدمة

1. الشبكات العصبية الالتفافية (CNNs)

تعتبر CNNs من أهم النماذج المستخدمة في رؤية الكمبيوتر، حيث تتألف من طبقات التلافيف والطبقات التجميعية التي تساعد في استخراج الميزات من الصور. من أشهر نماذج CNN:

- **LeNet**: نموذج مبسط يستخدم في التعرف على الأرقام المكتوبة يدويًا.

- **AlexNet**: قدم ثورة في تصنيف الصور بفضل عمقه وقدرته على التعلم من كميات ضخمة من البيانات.

- **ResNet**: استخدم مفهوم المسارات المتبقية لتدريب شبكات عميقة جداً دون مواجهة مشكلة تلاشي التدرجات.

2. الشبكات العصبية التكرارية (RNNs) والشبكات العصبية الطويلة والقصيرة المدى (LSTMs)
تستخدم هذه الشبكات لمعالجة البيانات التسلسلية مثل الفيديوهات، حيث تساعد في فهم السياق الزمني والتغيرات عبر الزمن.

3. شبكات التحويل (Transformers)

بدأت تستخدم في رؤية الكمبيوتر بفضل قدرتها على معالجة العلاقات البعيدة في البيانات وتحسين الأداء في مهام مثل الترجمة البصرية والتعرف على الأجسام.

التطبيقات العملية

السيارات ذاتية القيادة

تستخدم رؤية الكمبيوتر والتعلم العميق لتحليل البيئة المحيطة، بما في ذلك التعرف على العلامات المرورية، المشاة، والسيارات الأخرى، مما يمكن السيارة من اتخاذ قرارات القيادة الآمنة.

الرعاية الصحية

تساعد في تحليل الصور الطبية للكشف عن الأمراض مثل السرطان، وتحسين دقة التشخيص وتوفير الوقت للأطباء.

الأمان والمراقبة

تستخدم في أنظمة المراقبة لتحليل الفيديوهات.

عدد من النظريات والمفاهيم العلمية التي تشكل الأساس لفهم رؤية الكمبيوتر والتعلم العميق:

(1) نظرية الشبكات العصبية الاصطناعية (Artificial Neural Networks – ANN)

الشبكات العصبية الاصطناعية تشكل حجر الزاوية في التعلم العميق. تعتمد هذه الشبكات على هيكل مستوحى من الدماغ البشري، حيث تتكون من وحدات معالجة تُعرف بالعُقد أو الخلايا العصبية. تتعلم هذه الشبكات عن طريق تعديل الأوزان المرتبطة بالوصلات بين العُقد باستخدام تقنيات تحسين، مثل الانحدار التدرجي.

- **نظرية يونيفيرسال أ بروكسيمايتر (Universal Approximation Theorem):** تنص هذه النظرية على أن الشبكة العصبية الاصطناعية ذات الطبقة المخفية الواحدة يمكن أن تقارب أي دالة مستمرة إلى مستوى دقة معين، بشرط أن تكون الشبكة كبيرة بما يكفي.
- **تحليل التدرج العكسي (Backpropagation Algorithm):** هي تقنية تُستخدم لتحديث أوزان الشبكة العصبية بناءً على الخطأ في النتيجة المتوقعة مقابل النتيجة الفعلية، وتقوم هذه التقنية بحساب الانحدار العكسي من مخرجات النموذج إلى مدخلاته باستخدام قاعدة السلسلة.

(2) الشبكات العصبية الالتفافية (Convolutional Neural Networks – CNNs)

الشبكات العصبية الالتفافية تُستخدم بشكل خاص في معالجة البيانات المرئية، مثل الصور والفيديو. تعتمد هذه الشبكات على مفهوم العمليات الالتفافية (convolution operations) لاستخراج ميزات محلية من الصور.

- **نظرية الالتفاف (Convolution Theory):** الالتفاف هو عملية رياضية تتضمن دمج إشارة (مثل صورة) مع مصفوفة تسمى مرشح أو نواة.
- بعض النظريات التي تم تطويرها بناءً على دراسات وأبحاث أكاديمية. هذه النظريات تستند إلى مجموعة من الأوراق العلمية التي قدمت تطورات هامة في المجال:

(1) نظرية شبكات التفاف الأعصاب الالتفافية (CNN Theory):

شبكات الأعصاب الالتفافية (CNNs) هي واحدة من أكثر النظريات تأثيراً في رؤية الكمبيوتر، تأتي النظرية الأساسية وراء CNN من أبحاث **Yann LeCun** وآخرين في التسعينيات، وخاصة مع تقديم نموذج **LeNet-5** المستخدم للتعرف على الأرقام المكتوبة يدوياً.

النظرية: تعتمد CNN على فكرة العمليات الالتفافية، حيث يتم تمرير النواة عبر الصورة لاستخراج الميزات المحلية. تقوم الطبقات العميقة بالتعرف على الميزات الأكثر تعقيداً مثل الأشكال أو الأنماط في الصورة.

المراجع:

Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, 1998.

(2) نظرية "التعلم التمثيلي" (Representation Learning Theory) :

تدور هذه النظرية حول قدرة التعلم العميق على اكتساب ميزات تمثيلية للبيانات بشكل تلقائي في السابق، كان استخراج الميزات يتطلب تدخلاً بشرياً، لكن الشبكات العصبية العميقة قادرة على التعلم التلقائي للتمثيلات التي تحسن أداء المهام المختلفة مثل تصنيف الصور أو التعرف على الكلام.

النظرية: تنص على أن التعلم العميق يمكن أن يكتسب تمثيلات هرمية متعددة المستويات للبيانات، حيث تكون كل طبقة في الشبكة مسؤولة عن استخراج تمثيلات أكثر تجريداً وتعقيداً.

المراجع:

Bengio, Yoshua, et al. "Learning deep architectures for AI." *Foundations and trends in Machine Learning* 2.1 (2009): 1–127.

(3) نظرية "تحليل المكونات الأساسية العميقة – (Deep Principal Component Analysis) (PCA)

تعتمد هذه النظرية على تطبيق خوارزميات التحليل الإحصائي التقليدية مثل تحليل المكونات الأساسية (PCA) على الشبكات العصبية العميقة، تستخدم هذه الطريقة لتقليل أبعاد البيانات من خلال إسقاط البيانات على أساسات جديدة تسمى "المكونات الأساسية"، التي تمثل الأبعاد الأكثر تمييزاً في البيانات.

النظرية: يمكن للشبكات العصبية العميقة تقليد أداء PCA العادي، لكن على مستويات متعددة من التجريد. كل طبقة تعكس تحليلاً لمكونات أكثر تجريداً من البيانات.

المراجع:

Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504–507.

(4) نظرية "الانتقال المعرفي" (Transfer Learning Theory) :

تستند هذه النظرية إلى فكرة أن المعرفة المكتسبة من تعلم مهمة معينة يمكن استخدامها لتحسين الأداء في مهمة أخرى، حيث تم تقديم هذه النظرية بناءً على الأبحاث التي أظهرت أن نماذج التعلم العميق المدربة

على كميات ضخمة من البيانات (مثل شبكات CNN المدربة على مجموعة بيانات ImageNet) يمكن إعادة استخدامها لأغراض مختلفة بعد تعديلات طفيفة.

النظرية: النماذج المدربة مسبقاً تستطيع نقل المعرفة التي اكتسبتها في بيئة معينة إلى مهام جديدة، مما يقلل الحاجة إلى كميات ضخمة من البيانات أو تدريب طويل.

المراجع:

Yosinski, Jason, et al. "How transferable are features in deep neural networks?"

Advances in neural information processing systems. 2014.

5) نظرية "تلاشي التدرجات" (Vanishing Gradient Theory)

هذه النظرية تتناول مشكلة تواجه الشبكات العصبية العميقة عند تدريبها، تم اكتشاف أن استخدام وظائف تفعيل تقليدية مثل السيجمويد يتسبب في تقليص القيم المنتشرة خلال طبقات الشبكة، مما يجعل من الصعب تدريب الطبقات العميقة بفعالية.

النظرية: عندما تصبح التدرجات صغيرة جداً، تفشل الشبكة في تحسين الأوزان بشكل صحيح خلال عملية التدريب، مما يؤدي إلى تدهور الأداء. أدى هذا الاكتشاف إلى تطوير تقنيات مثل استخدام وظيفة تفعيل

ReLU.

المراجع:

Hochreiter, Sepp. "The vanishing gradient problem during learning recurrent neural nets and problem solutions." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 6.02 (1998): 107–116.

6) نظرية "تقدير الاحتمالات العميق" (Deep Probabilistic Modeling)

تستند هذه النظرية إلى فكرة أن الشبكات العصبية يمكن أن تعمل كنماذج احتمالية للتنبؤ بالنتائج بناءً على بيانات المدخلات، تعمل هذه النماذج بشكل كبير على تحسين المهام التي تتطلب التعامل مع عدم اليقين، مثل التعرف على الصور أو الكلام.

النظرية: يمكن لنماذج التعلم العميق تقدير الاحتمالات المتعلقة بمختلف النتائج الممكنة، مما يجعلها مناسبة للمهام التي تتطلب التعامل مع ضوضاء أو بيانات غير مؤكدة.

المراجع:

Bishop, C. M. "Pattern recognition and machine learning." *Springer*, 2006.

حيث ان البحوث العلمية في مجال رؤية الكمبيوتر والتعلم العميق تعتمد بشكل كبير على نظريات رياضية وإحصائية معقدة تعزز من فهم الآلات للبيانات البصرية. هذه النظريات تم اختبارها وتطويرها عبر عقود من الأبحاث الأكاديمية والتجريبية، وساهمت بشكل كبير في تقدم الذكاء الاصطناعي وتقنيات تعلم الآلة.