# 2-1 مقدمة عن الرؤية بالحاسوب (Computer Vision) :

ظلت رؤية الكمبيوتر موجودة منذ أكثر من 50 عامًا ، ولكن في الآونة الأخيرة ، نشهد زيادة كبيرة في الاهتمام بكيفية "رؤية" الآلات وكيفية استخدام رؤية الكمبيوتر لإنشاء منتجات للمستهلكين والشركات. بعض الأمثلة على هذه التطبيقات هي: Amazon Go و Google Lens و Autonomous Vehicles و Face Recognition.

العامل الرئيسي وراء كل هذا هو رؤية الكمبيوتر. بعبارات أبسط ، تمثل Computer Vision مجالا واسعا من الذكاء الاصطناعي الذي يعلم الآلات أن ترى. هدفها هو استخراج المعنى من البكسلات.

من وجهة نظر العلوم البيولوجية ، تتمثل أهدافها في التوصل إلى نماذج حسابية للنظام البصري البشري. من وجهة النظر الهندسية ، تهدف رؤية الكمبيوتر إلى بناء أنظمة مستقلة يمكنها أداء بعض المهام التي يمكن أن يؤديها النظام البصري البشري (بل وتجاوزها في كثير من الحالات).

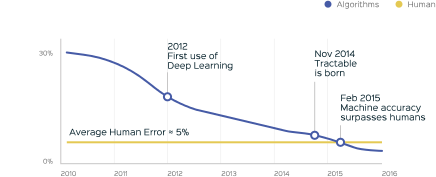
**2-2 لمحة تاريخية :**

في صيف العام 1966 ، بدأ سيمور بابيرت ومارفين مينسكي من مجموعة MIT Artificial Intelligence مشروعًا بعنوان مشروع الرؤية الصيفية. كان الهدف من المشروع هو بناء نظام يمكنه تحليل المشهد وتحديد الكائنات في المشهد. لذا فإن المجال الواسع المحير المتمثل في رؤية الكمبيوتر الذي لا يزال الباحثون وعمالقة التقنية يحاولون فك تشفيره ، كان يُعتقد أولاً أنه بسيط بما فيه الكفاية لمشروع صيفي لمرحلة ما قبل التخرج من قبل نفس الأشخاص الذين كانوا رواد في مجال الذكاء الاصطناعي.

في سبعينيات القرن الماضي ، أخذ ديفيد مار ، وهو عالم الأعصاب في معهد ماساتشوستس للتكنولوجيا ، الذي أخذ أفكارًا من دراسات المخيخ والحصين ، اللبنات الأساسية لرؤية الكمبيوتر الحديثة ، وبالتالي يعرف باسم والد رؤية الكمبيوتر الحديثة. تتوج غالبية أفكاره في الكتاب الذي يحمل عنوان VISION.

**2-3 الرؤية العميقة (Deep Vision) :**

لقد بدأ التعلم العميق (Deep Learning) منذ عام 2012. التعلم العميق هو مجموعة فرعية من التعلم الآلي (machine learning) حيث تتعلم الشبكات العصبية الاصطناعية ، الخوارزميات المستوحاة من الدماغ البشري ، من كميات كبيرة من البيانات. تشغيل أنظمة التوصية ، وتحديد الأصدقاء في الصور ووضع علامات عليهم ، وترجمة صوتك إلى نص ، وترجمة النص إلى لغات مختلفة ، وقد حوّل التعلم العميق رؤية الكمبيوتر إلى أداء أكثر تفوقاً.



معدل الخطأ في تصنيف الصور مع مرور الوقت ، انخفاض حاد بعد إدخال التعلم العميق. المصدر - tractable.ai

**2-4 استخدامات رؤية الحاسوب :**

* الهواتف الذكية: أكواد QR ، والتصوير الحسابي (Android Lens Blur ، iPhone Portrait Mode) ، إنشاء بانوراما (Google Photo Spheres) ، اكتشاف الوجه ، اكتشاف التعبير (الابتسامة) ، فلاتر Snapchat (تتبع الوجه) ، Google Lens ، Night Sight (Pixel).
* الويب: البحث عن الصور ، صور Google (التعرف على الوجوه ، التعرف على الأشياء ، التعرف على المشهد ، تحديد الموقع الجغرافي من الرؤية) ، Facebook (شرح الصورة) ، خرائط Google للتصوير الجوي (خياطة الصورة) ، YouTube (تصنيف المحتوى).
* VR / AR: التتبع الخارجي (HTC VIVE) ، التتبع الداخلي (التعريب والتخطيط المتزامنان ، HoloLens) ، انسداد الكائنات (تقدير العمق الكثيف).
* التصوير الطبي: إعادة بناء CAT / التصوير بالرنين المغناطيسي ، والتشخيص بمساعدة ، وعلم الأمراض التلقائي ، و connectomics ، والجراحة الموجهة بالذكاء الاصطناعي (AI-guided surgery).
* الوسائط: التأثيرات المرئية للفيلم ، والتلفزيون (إعادة الإعمار) ، وإعادة الرياضة الافتراضية (إعادة الإعمار) ، والتعديلات التلقائية المستندة إلى دلالات (إعادة الإعمار ، والاعتراف).
* التأمين: أتمتة المطالبات ، تحليل الأضرار ، فحص الممتلكات.

**2-5 التحديات :**

* حتى بعد نشر قدر كبير من العمل ، لا يتم حل رؤية الكمبيوتر. إنه يعمل فقط في ظل بعض القيود. أحد الأسباب الرئيسية لهذه الصعوبة هو أن النظام البصري البشري هو ببساطة جيد جدًا للعديد من المهام مثل التعرف على الوجوه. يمكن للإنسان التعرف على الوجوه تحت جميع أنواع الاختلافات في الإضاءة ، وجهة النظر ، التعبير ، وما إلى ذلك الذي يعاني منه الكمبيوتر في مثل هذه الحالات.
* الخصوصية والأخلاق - أثناء استخدام المراقبة ، تشتمل أحدث قطاعات التأمين على تعديل الأقساط والسياسات من خلال مراقبة سلوك القيادة ، لكن من ناحية أخرى ، تشكل أنظمة المراقبة التي تعمل بالطاقة Vision مخاطر كبيرة على الخصوصية والأخلاق. كمثال ، نرى الصين ، باستخدام التعرف على الوجه لتتبع الأقليات العرقية. في الآونة الأخيرة ، أصبحت سان فرانسيسكو أول مدينة أمريكية تمنع استخدام التعرف على الوجه من قبل حكومتها.
* عدم وجود تفسير - لا تزال الخوارزميات الحديثة القائمة على الشبكة العصبية صندوقًا أسودًا إلى حد ما. لذلك عندما يصنف نموذج صورة كسيارة ، فإننا لا نعرف في الواقع سبب قيامها بذلك. تعد "قابلية التفسير" شرطا أساسيا في العديد من المجالات مثل التأمين والقيادة الذاتية التي تفتقد حاليا في هذه الخوارزميات.
* Deep Fakes -. باستخدام تقنيات التعليم العميق ، يمكن الآن لأي شخص لديه بيانات GPU وبيانات تدريب قوية إنشاء صورة مزيفة أو مقاطع فيديو يمكن تصديقها باستخدام DeepFakes. هذه المشكلة خطيرة للغاية ، حيث يعمل البنتاغون ، من خلال وكالة مشاريع الأبحاث المتقدمة للدفاع (DARPA) ، مع العديد من أكبر المؤسسات البحثية في البلاد لمعالجة DeepFakes.
* هجمات الخصوم (Adversarial attacks) - أمثلة الخصم هي مدخلات لنماذج التعلم الآلي التي صممها المهاجم عمدا لتتسبب في ارتكاب أخطاء ؛ مثل الأوهام البصرية للآلات.

**2-6 مستقبل رؤية الكمبيوتر :**

وفقًا لتقرير ، بلغت قيمة سوق رؤية الكمبيوتر 2.37 مليار دولار أمريكي في عام 2017 ، ومن المتوقع أن تصل إلى 25.32 مليار دولار أمريكي بحلول عام 2023 ، بمعدل سنوي مركب نسبته 47.54٪.

يشهد العالم تحولًا رقميًا عميقًا ، خاصةً الهند التي لا تظهر عليها أي علامات على التباطؤ. متوسط استهلاك البيانات الشهري لـ Jio وحده هو 10.8 جيجابايت. وفقًا لهذا التقرير ، كل دقيقة:

* شاهد المستخدمين 4،146،600 مقطع فيديو على YouTube.
* ينشر مستخدمو Instagram 46،740 صورة.
* يشارك مستخدمو Snapchat 527،760 صورة.

التي تعطي جميعها مجموعة كبيرة من الفرص لرؤية الكمبيوتر لتجد أنماطًا وتفهمها.

حتى مع كل التطورات الرائعة ، يحتاج الذكاء الاصطناعى ومجال رؤية الكمبيوتر على وجه التحديد إلى معالجة المشاكل المرتبطة به حاليًا مثل التحيز ، وخطر عدم الوعي وعدم وجود تفسير. لمعالجة مثل هذه المشكلات ، بدأت شركات مثل Ping An في اتخاذ خطوات صغيرة ، باستخدام Symbolic AI ، وهو شكل مبكر من الذكاء الاصطناعي ، في خوارزميات الذكاء الاصطناعى الحديثة لإعطائها القدرة لتوضيح قرارها ، ولكن لا يزال هناك الكثير للقيام به.