كيف تعمل خوارزميات التعرف على الوجوه جيدًا مع ملايين الغرباء؟



تتضمن مجموعة بيانات المليون وجه MegaFace مليون صورة تمثّل أكثر من 690,000 من الأشخاص المختلفين. وتعد هذه المجموعة أوَّل مرجع يختبر خوارزميّات الكشف عن الوجوه ضمن معدّل مليون شخص.

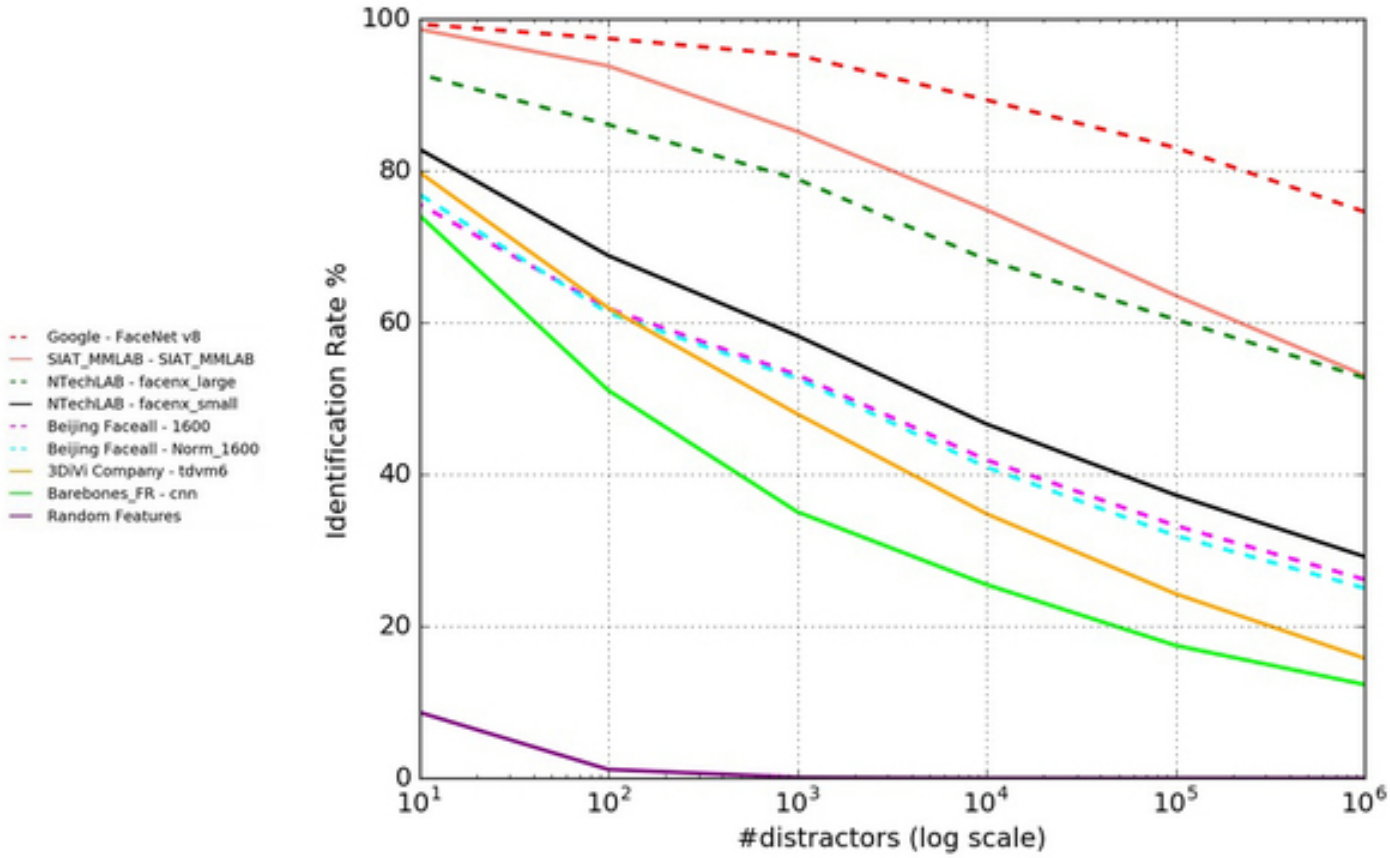
في السنوات القليلة الماضية، أعلنت مجموعات متعددة أنَّ أنظمة التعرّف على الوجوه الخاصة بهم حقّقت معدلات دقّة شبه مثالية، وبأداء أفضل من البشر، في اختيار نفس الوجه ضمن حشود كبيرة من الناس.

إلا أن هذه الاختبارات أُجريت على مجموعة بيانات مؤلفة فقط من 13,000 صورة، وهو عدد أقل من الأشخاص الذين يشاهدون لعبة كرة قدم أمريكية عاديّة. فما الذي قد يحدث لو أنّ عددهم تضاعف ليصل عدد سكان مدينة أمريكية.

أجاب باحثون في جامعة واشنطن على هذا السؤال من خلال تحدي المليون وجه، وهي أوّل مسابقة عالمية هدفها تقييم التحسّن في أداء خوارزميات التعرّف على الوجوه بمعدّل مليون شخص. وقد عانت كل الخوارزميات من موضوع الدقة، بعد مواجهتها بمزيد من الصور، لكن بعضها كان أفضل من بعضها الآخر.

يقول أستاذ مساعد في علوم الحاسب في جامعة واشنطن والباحث الرئيسي في المشروع، إيرا كاميلماتشر شلايزرمان **Ira Kemelmacher-Shlizerman**: "نريد أن نفحص التعرّف على الوجوه على مستوى الكوكب كله، بهدف استخدام ذلك في تطبيقات حقيقية –حيث إنّ الفحص بهذه الطريقة يسمح باكتشاف نقاط نجاح وفشل الخوارزميّة". ويتابع قائلًا: "لا نستطيع فحص الخوارزميّات على مجموعات صغيرة من البشر، ثم نقول أنّ الخوارزميّة تعمل بنجاح".

في البداية، أنشأ فريق جامعة واشنطن مجموعة بيانات مؤلفة من مليون صورة مجانية من مختلف أنحاء العالم تمثل بمجموعها 690,572 شخصًا مختلفًا. من ثم سمحوا للفرق المشاركة بتحميل هذه الصور وفحص أداء الخوارزميات الخاصة بهم في التعرف على الوجوه عندما يكون عليها التمييز بين مليون حالة تطابق محتملة.



خوارزميّات التعرّف على الوجوه التي قامت بأداء جيّد ضمن 10,000 صورة مختلفة، جميعها قد واجهت انخفاضًا في الدقّة عندما تمّت مواجهتها بمليون صورة، إلّا أنّ بعضها قام بأداء أفضل من بعضها الآخر. حقوق الصورة: جامعة واشنطن University of Washimgton.

وقد كان أداء جوجل فيس نت **FaceNet** هو الأفضل في هذا الاختبار، مع انحدار بالدقّة، التي كانت شبه مثاليّة عند استخدام خوارزميتهم مع عدد أقل من الصور، إلى 75% عند وضعها ضمن اختبار المليون صورة. وقد جاء فريق روسي من **N_TechLab** أولاً على الفرق الأخرى التي استخدمت مجموعة بيانات أخرى، إلا أن دقته انحدرت إلى 73%..

وعلى النقيض من ذلك، فإن معدلات دقة الخوارزميات الأخرى، التي كان أداؤها جيداً (أكثر من 95%) مع عدد قليل من الصور، انحدرت بنسبة أكبر كثيراً بنحو 33% عندما واجهت مهمة أصعب.

وقد نُوقشت النتائج الأوليّة في ورقة بحث علمي قدّمت للنشر في مؤتمر **IEEE** المختص في الرؤية الحاسوبيّة وتمييز الأنماط (**CVPR 2016**) في 30 حزيران/ يونيو، وتُحدث النتائج المستمرة في الموقع الخاص بالمشروع. وهناك أكثر من 300 مجموعة بحثية تعمل على مشروع ميجا فيس.



سلّط اختبار المليون وجه الضوء على المشاكل التي تعتري خوارزميّات التعرّف على الوجوه والتي لم يتم حلها بشكل كامل بعد، مثل تمييز نفس الشخص عند أعمار مختلفة أو التعرّف على شخص واحد بعدّة وضعيات. حقوق الصورة: جامعة واشنطن.

وقد اختبر تحدي ميجا فيس الخوارزميات من حيث التحقّق، أو كيف سيكون أداؤها من حيث تمييز ما إذا كانت الصورتان للشخص نفسه أم لا. على سبيل المثال، هكذا تعمل ميزة الأمان في أجهزة الـ **iPhone** على تمييز وجه المستخدم ومن ثم تقرر فتح القفل بدلًا من طلب إدخال كلمة المرور منه.

يقول المدير المساعد في مختبر الرسوم والتصوير في جامعة واشنطن (**GRAIL**.)، شلايزرمان: "ما الذي سيحصل إذا أضعت هاتفك في محطة قطارات أمستردام وحاول شخص ما سرقته؟". ويتابع: "أودُّ أن يميزني هاتفي بشكل صحيح من بين مليون شخص – أو 7 مليار – وليس فقط 10,000 شخص أو نحوهم".

لقد سلّط اختبار المليون وجه الضوء على المشاكل التي تعتري خوارزميّات التعرّف على الوجوه، والتي لم تُحل بشكل كامل بعد، على سبيل المثال تمييز نفس الشخص في أعمار مختلفة أو التعرّف على شخص واحد في عدّة وضعيات.

كما فُحصت الخوارزميات أيضا من حيث تحديد الهويات، أو من حيث مدى دقّتها في العثور على تطابق بين صورة شخص ما مع صورة مختلفة لنفس الشخص، موضوعة ضمن مجموعة من مليون صورة مختلفة. وهذا ما حصل، عندما تكون لدى دائرة مكافحة الجريمة صورة واحدة لمجرم أو مشتبه به، فإنّ الخوارزمية تُمشط الصور التي ألتقطت من محطة المترو أو المطار لرؤية المجرم عند محاولته الهروب.

يقول شلايزرمان: "بوسعك أن ترى أين تكمن المشاكل الصعبة، فتمييز الناس على فئات عمرية مختلفة، هي مشكلة لم تحل بعد. وكذلك الأمر مع التعرّف على الناس من خلال أشباههم ومقارنة الناس من أكثر من موقع، من مثل المنظر الجانبي والأمامي". وتحلّل الورقة العلمية أيضًا الأعمار وثبات التموضّع بتمييز الأوجه عند مقارنتها بتدرج معين.

بشكل عام، فإن الخوارزميات التي تدرّبت على إيجاد مقارنات صحيحة من خلال العمل على مجموعة كبيرة من الصور، تغلبت على مثيلاتها التي تعاملت مع عدد أقل من الصور. لكن خوارزمية **SIAT MMLab**، التي طوّرتها فريق بحث في الصين، والتي تدرّبت على عدد أقل من الصور، صُنفت ضمن الخوارزميات الجيدة.

ولازال ميجا فيس مستمر بالمنافسة وتقبّل النتائج.

تتضمّن الخطوة التالية لفرق البحث تجميع نصف مليون شخصية ـ كل منها مزود بعدة صور ـ لتجميعها في مجموعة بيانات سوف تستخدم لتدريب خوارزميات التعرف على الأوجه. سوف يساعد هذا على رفع مستوى التمييز بين الخوارزميات المشاركة ومعرفة من منها سيكون ذا أداء جيد مع نفس الكمية الكبيرة من الصور. ولأن معظم الباحثين ليس بوسعهم الوصول لمجموعات الصور الموجودة، من مثل الموجودة لدى جوجل أو فيس بوك، فستُنشر مجموعة الصور الخاصة بالفحص مع نهاية الصيف.

يقول طالب الماجستير في علوم وهندسة الحاسوب في جامعة واشنطن الذي يعمل على مجموعة بيانات لتدريب الخوارزميات، آرون نيك **Aaron Nech**: "لدى خوارزميات الشبكات العصبيّة ذات التعلّم العميق، والتي تستخدم أحدث التقنيّات، ملايين من المعاملات لتتعلّمها، كما أنها تتطلّب عددًا ضخمًا من الأمثلة من أجل توليف هذه الشبكات بشكل جيّد". ويضيف قائلًا: "بشكل مختلف عن البشر، هذه النماذج تشكّل في البداية لائحة فارغة. ووجود اختلاف وتباين في البيانات، من مثل أدلة التمييز المعقّدة الموجودة لدى أكثر من نصف مليون شخص مختلف، من الممكن أن يرفع من أداء الخوارزمية، من خلال تزويدها بأمثلة لحالات لم تُفحص بعد".

- التاريخ: 2018-09-18
- التصنيف: كيف تعمل الأشياء؟

#التعلم_العميق #الشبكات_العصبية #خوارزميات #ذكاء_اصطناعي

المصادر

- phys.org

المساهمون

- ترجمة
  - أمجد هواش
- مُراجعة
  - محمد اسماعيل باشا
- تحرير
  - طارق نصر
  - بنان محمود جوابره
- تصميم
  - علي كاظم
- نشر
  - بنان محمود جوابره