

يوتيوب تضيف الوصف التلقائي للتأثيرات الصوتية



الجمعة 24 مارس 2017 10:03 م

أطلقت منصة مشاركة مقاطع الفيديو [يوتيوب](#) المملوكة لشركة جوجل نظامها الجديد للتسميات التوضيحية للتأثيرات الصوتية الذي يمكنه أن يحدد ويصف بشكل تلقائي التأثيرات الصوتية المتمثلة حالياً بالموسيقى والصوت والضحك الموجودة ضمن مقاطع الفيديو، ووعدت الشركة بإضافة المزيد من القدرات خلال المراحل القادمة مع تدريبها لخوارزميات الكشف []
واستخدمت يوتيوب منذ عام 2009 حتى الآن الخوارزميات لتوضيح الكلام الموجود ضمن الفيديو بشكل تلقائي، وحاولت على مدى السنوات الثمانية السابقة جعل المليارات من مقاطع الفيديو الموجودة لديها أكثر سهولة من ناحية الفهم بالنسبة للمستخدمين الذين يعانون من الصمم أو ضعف السمع []
وقد تحسنت الميزة بشكل ملحوظ مع مرور الوقت بعكس ما كانت عليه في بداية الأمر، وبحسب مدونة نشرتها جوجل على موقعها للمطورين فإن الميزة أصبحت أقرب أكثر من أي وقت مضى إلى معدلات الخطأ البشري []
وتعتبر الإضافة صغيرة إلا انها مهمة بالنسبة لميزات إمكانية الوصول لمنصة مشاركة مقاطع الفيديو على الإنترنت، وتعد إمكانية توضيح التأثيرات الصوتية مهمة صعبة، ووفقاً لمهندسي جوجل فإن هذه المهمة لم يكن من الممكن تحقيقها إلا عبر التطوير التي حدثت على مدى العامين الماضيين لتقنيات التعلم الآلي []
وبحسب الشركة فإن المشكلة لا تتعلق بشكل كبير بإمكانيات الحواسيب في الكشف عن الأشياء وتصنيفها، حيث أن منتجات مثل جوجل للصور Photos على سبيل المثال تمتلك قدرات كشف عن الصور، إلا أن المشكلة تتعلق بعدم وجود قاعدة بيانات هامة للمؤثرات الصوتية يمكن استخدامها عند تدريب الشبكة العصبية القادرة على الكشف والتحديد []
وقرر الفريق التركيز على الموسيقى والصوت والضحك بشكل أولي لأنها تضيف معنى مهم للحوار الموجود ضمن الفيديو بالنسبة للأشخاص الذين يعانون من الصمم أو من صعوبة السمع، وقد أصبح لدى يوتيوب خوارزمية ذكاء صناعي مدربة مخصصة للتأثيرات الصوتية، وذلك بعد معالجة آلاف الساعات من مقاطع الفيديو []
وأشارت الشركة إلى ان التحديات المستقبلية قد تتضمن إمكانية إضافة فئات وأنماط صوتية شائعة أخرى مثل الرنين والنباح والنقر، والتي قد تمثل مشاكل معينة على سبيل المثال مع صوت الرنين فإن الشركة بحاجة إلى تكون قادرة على تحديد فيما إذا كان هذا الرنين رنين المنبه أو رنين جرس الباب أو رنين الهاتف []