

أداة ذكية جديدة من "ميتا" تنظف التسجيلات الصوتية المعقدة

20 ديسمبر، 2025



أعلنت شركة ميتا عن إطلاق أداة ذكاء اصطناعي جديدة مفتوحة المصدر باسم SAM Audio، قادرة على فصل الأصوات وتنقية التسجيلات المزدحمة بالضوضاء بمجرد إدخال وصف نصي للصوت المطلوب، دون الحاجة إلى برامج تحرير صوتية معقدة.

تتيح الأداة للمستخدمين استخراج أصوات محددة مثل الصوت البشري، الآلات الموسيقية، أو ضوضاء الخلفية من تسجيل واحد مزدحم، ما قد يحدث تحولاً كبيراً في مجالات البودكاست، الإنتاج الموسيقي، صناعة الأفلام، وتسهيل الوصول لذوي الإعاقة. وتعتمد SAM Audio على نموذج متعدد الوسائط يدعم ثلاثة أنواع من الأوامر:

وصف نصي للصوت المراد عزله.

اختيار بصري لشخص أو عنصر داخل الفيديو لتحديد مصدر الصوت.

تحديد زمني للحظة ظهور الصوت في التسجيل.

يمكن استخدام هذه الطرق بشكل منفصل أو مجتمعة لتحقيق دقة أعلى في فصل الأصوات. ويستند النموذج إلى محرك Perception Encoder

Audiovisual من ميتا، الذي يحلل الأصوات قبل فصلها بدقة عن بقية المقطع، وفق تقرير موقع "DigitalTrends".

كما كشفت ميتا عن معيار تقييم جديد باسم SAM Audio-Bench لقياس أداء نماذج فصل الصوت، إلى جانب أداة SAM Audio Judge لتقييم جودة الصوت الناتج من منظور المستمعين، حتى دون وجود ملف مرجعي. وأظهرت النتائج تفوق النموذج عند الجمع بين أكثر من نوع من الأوامر، مع قدرته على معالجة الصوت بسرعة تفوق الزمن الحقيقي، حتى على نطاق واسع.

تشمل الاستخدامات العملية للأداة:

إزالة ضوضاء المرور من تسجيلات البودكاست.

عزل الصوت البشري عن التسجيلات الموسيقية.

حذف أصوات غير مرغوبة مثل نباح الكلاب من التسجيلات المنزلية.

ومع ذلك، تواجه SAM Audio بعض القيود الحالية، مثل عدم دعم الأوامر الصوتية، وصعوبة فصل الأصوات المتشابهة أو المتداخلة دون توجيه. ويأتي هذا الإطلاق ضمن توجه ميتا لتطوير نماذج ذكاء اصطناعي تفهم الصوت والسياق والتفاعل البشري، بما في ذلك تحسين وضوح الصوت في نظاراتها الذكية، وتطوير نظارات واقع مختلط متقدمة بحلول 2027، وبناء مساعد محادثة ذكي قد ينافس شات جي بي تي.

الأداة أصبحت متاحة الآن عبر Segment Anything Playground، مما يعزز جهود ميتا في إتاحة تقنيات الذكاء الاصطناعي المتقدمة للمطورين وصناع المحتوى حول العالم.