

كيفاش NuminaMath ربح الجائزة ديال الذكاء الاصطناعي فالرياضيات (AIMO)

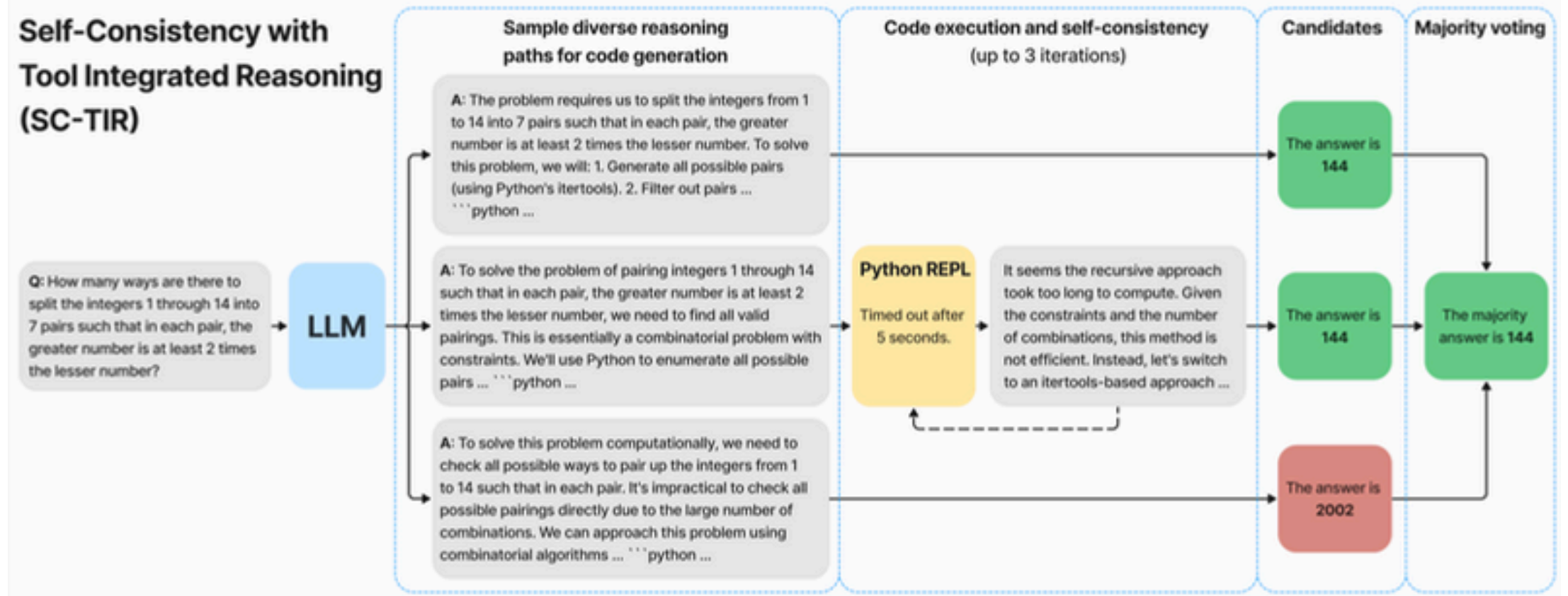
نومينا Numina و Hugging Face تعاونو باش يطورو NuminaMath 7B TIR، واحد الموديل ديال Natural Language اللي تميز فالأولمبياد ديال الذكاء الاصطناعي فالرياضيات. الأداء ديال الموديل، لي حلّ تقريبا 60% من المسائل دالمات، كيبين القدرة دبالو على التعامل مع الأسئلة الرياضية لي صعبة شوية. هاد الإنجاز كيبين على الإمكانيات ديال Ai باش ادير ثورة فحل المسائل الرياضية لي فالمستوى نتاع الالمبياد.

NuminaMath 7B TIR تطور عبر من بعد fine-tuning لي داز من جوج ديال المراحل. فاللول، الموديل الاول تدرب على واحد dataset ديال مسائل رياضية باش ييني أساس ديال المعرفة الرياضية و reasoning. فالتاني، تحسن هاد model باستعمال واحد synthetic dataset اللي ركز على tool-integrated reasoning (TIR). المسائل الرياضية تقسمات لسلسلة ديال تفسيرات منطقية لبعث التمارين، كود نتاع Python، والنتائج ديالها. هاد الطريقة كتستعمل GPT-4 باش تلقى حلول اللي فيها كود Python.

- الموديل اللي تخدم فهاد المسابقة استعمل الطريقة نتاع TIR (شوف الصورة لتحت):
- بالنسبة لكل مشكل رياضي، inputs تعاودو بزاف دمرات باش نخلقو batch فاللول ديال prompts. هاد الشي كيحدد شحال دالحلول اللي غادي اعتبارها فالتصويت بالأغلبية فاللخر.
- من بعد كنتجو واحد set ديال completions كنجمعوها و نحصلو على block كامل ديال كود Python لكل حل محتمل.
- كل block ديال Python كيتم execution دبالو، وكيجمعو النتائج.
- هاد العملية تكررات بزاف دمرات باش نخلقو batch ديال generations مع عدد من محاولات، هاد الشي سمح للموديل يصحح أخطاء الكود بالاعتماد على feedback.
- ف اللخر أي حلول اللي جابت نتائج غير منطقية كتصفى، ومن بعد كيدار التصويت بالأغلبية majority voting باش نختارو الجواب النهائي.

References:

- <https://huggingface.co/blog/winning-aimo-progress-prize>
- <https://huggingface.co/spaces/AI-MO/math-olympiad-solver>



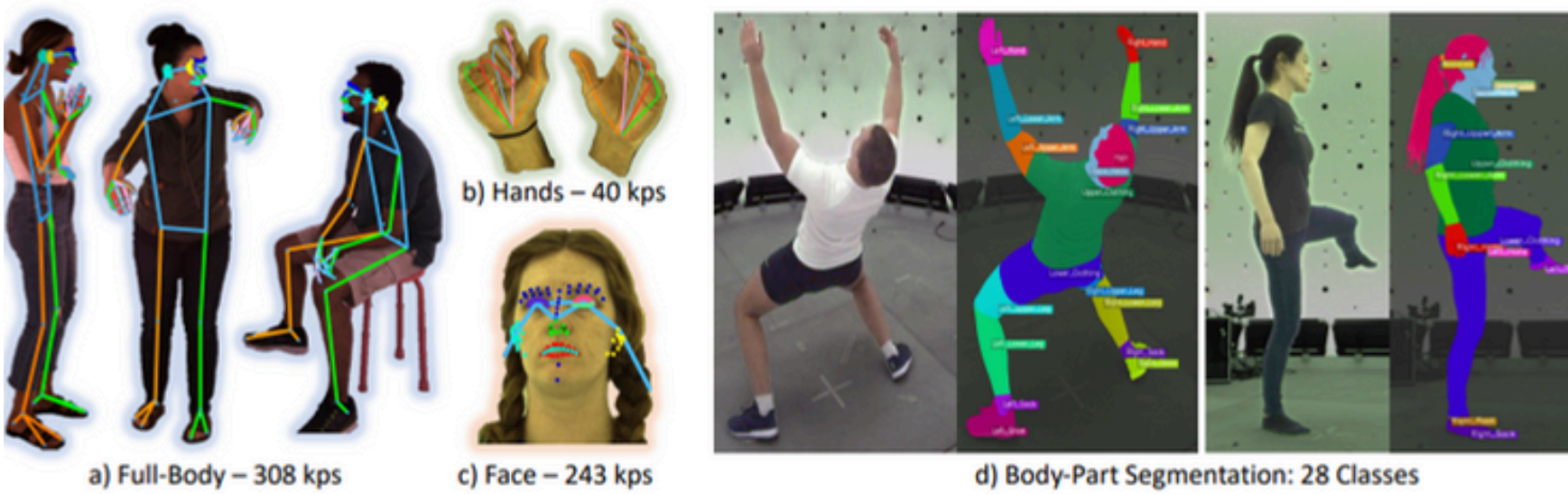
NuminaMath 7B TIR هو أداة داو صالحة لحل المسائل الرياضية لي معقدة شوية، ولكن يمكن تواجه صعوبات مع المسائل الأكثر تعقيدًا اللي كتجاوز مستوى 12 level (AMC) American Mathematics Competitions، خصوصًا فالدومين دالهندسة Geometry. فالمجمل، الإصدار ديال NuminaMath 7B TIR كيغطي واحد الوسيلة للناس اللي كيتعاملو مع تحديات رياضية عالية المستوى، وكيبين الإمكانيات ديال الذكاء الاصطناعي باش ادير ثورة فحل المسائل الرياضية.

إعادة تعريف الذكاء الاصطناعي اللي موجه نحو الإنسان باش نفهمو الرؤية ديال الجيل الجديد

Model	#Params	FLOPs	Hidden size	Layers	Heads	Batch size
Sapiens-0.3B	0.336 B	1.242 T	1024	24	16	98,304
Sapiens-0.6B	0.664 B	2.583 T	1280	32	16	65,536
Sapiens-1B	1.169 B	4.647 T	1536	40	24	40,960
Sapiens-2B	2.163 B	8.709 T	1920	48	32	20,480

مواصفات مشفر Sapiens للتدريب المسبق على مجموعة بيانات Human-300M

- تتميز Sapiens بقدرات استدلال عالية الدقة (1K) وقابلية للتوسع، مع نماذج كتراوح بين 0.3 حتى 2 مليار من المعاملات، متفوقة بزاف على النماذج اللي كاينة دابا عبر مختلف المعايير اللي عندها علاقة بالإنسان.
- باستخدام مجموعة بيانات منسقة فيها 300 مليون صورة بشرية فبيئات طبيعية، كاتبان Sapiens عندها قدرة عالية على التعميم، حتى فالحالات اللي فيها بيانات معنونة محدودة أو مجموعات بيانات اصطناعية بالكامل. هاد الشي كيجعلها الخيار المثالي للتطبيقات اللي كتححتاج تمثيل وفهم دقيق للبشر.
- كيبان أداء النموذج فالمهام بحال تقدير الوضع، تقسيم أجزاء الجسم، تقدير العمق، وتوقع الأسطح العادية تحسن كبير، مع تحسينات توصل حتى 53.5% فالغلط الزاوي النسبي على المعايير بحال THuman2. وزيد على هاد الشي، Sapiens كتتفوق فالحالات الواقعية، فين كتقدر تتعامل مع الصور البشرية المعقدة والمتنوعة بدقة ما عمرها كانت من قبل.



Ground-truth annotations for 2D pose estimation and body-part segmentation.

References:

- <https://arxiv.org/pdf/2408.12569>

- فاش كنجمعو مابين قابلية التوسع وبدقة عالية والاستخدام الفعال للتدريب الذاتي، Sapiens Vision Transformer كاتستعد باش تولي مابين اهم التطبيقات اللي كتعطينا مابين رقمنة البشر حتى لنمذجة ثلاثية الأبعاد المتقدمة، وحتت معايير جديدة لداك الشي اللي يمكن نحققوه فروية الكمبيوتر اللي موجهة نحو الإنسان.

official partner



www.mathemaroc.com

math&maroc



كيف فاز NuminaMath بالجائزة الأولى في أولمبياد للذكاء الاصطناعي في الرياضيات (AIMO)؟

تعاونت شركتنا نوميना Numina و Hugging Face لتطوير NuminaMath 7B TIR، وهو نموذج لغوي تميز في أولمبياد الرياضيات للذكاء الاصطناعي. الأداء المثير للإعجاب لهذا النموذج، حيث تمكن من حل حوالي 60% من المسائل الرياضية المعقدة، يوضح قدرته على التعامل الفعال مع الاستفسارات الرياضية الصعبة. هذا الإنجاز يبرز الإمكانيات الهائلة للذكاء الاصطناعي في إحداث ثورة في حل المسائل الرياضية.

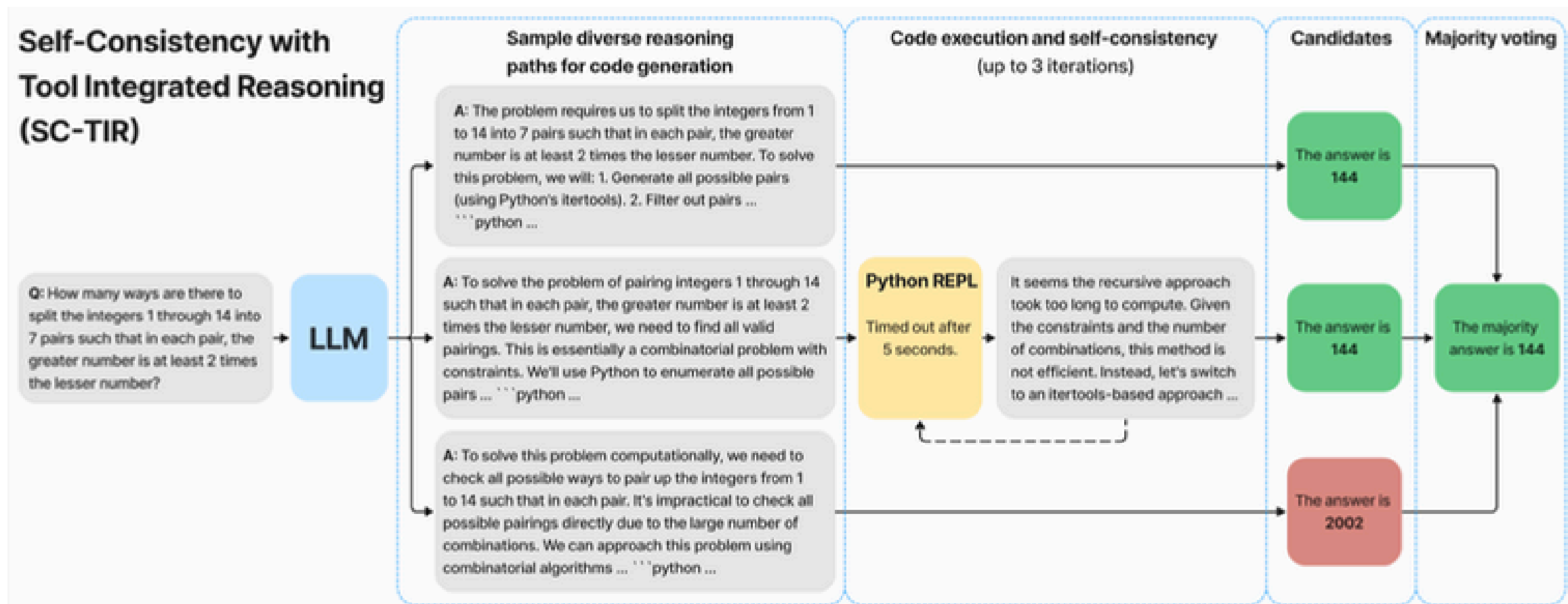
تم تطوير NuminaMath 7B TIR من خلال عملية finetuning تتكون من مرحلتين. أولاً، تم تدريب النموذج الأساسي على مجموعة بيانات واسعة من المسائل الرياضية لتأسيس قاعدة من المعرفة الرياضية. ثانياً، تم تحسينه باستخدام مجموعة synthetic dataset ركزت على tool-integrated reasoning (TIR). تم تقسيم المسائل الرياضية إلى سلسلة من التفسيرات المنطقية، وبرامج Python، والنتائج المقابلة لها. هذه المقاربة تستفيد من GPT-4 لتوليد حلول تتضمن كود Python قابل للتنفيذ.

استخدم النموذج المطبق نهج TIR (تحقق من الشكل أدناه):

- * لكل مشكلة، تم تكرار inputs عدة مرات لإنشاء دفعة أولية من batches. هذا حدد عدد الحلول المرشحة التي تم اعتبارها في التصويت بالأغلبية.
- * تم توليد مجموعة متنوعة من الإكملات completions حتى تم إنتاج block كاملة من كود Python لكل مرشح.
- * تم تنفيذ كل مجموعة Python، وتم دمج النتائج.
- * تكررت هذه العملية عدة مرات لإنشاء دفعة من generations مع محاولات متعددة، مما أتاح للنموذج تصحيح أخطاء الكود استناداً إلى feedback.
- * تمت تصفية أي مرشحين الذين أنتجوا مخرجات غير منطقية، ثم تطبيق التصويت بالأغلبية لاختيار الإجابة النهائية.

References:

- <https://huggingface.co/blog/winning-aimo-progress-prize>
- <https://huggingface.co/spaces/AI-MO/math-olympiad-solver>



NuminaMath 7B TIR هو أداة ذكاء اصطناعي قوية لحل المسائل الرياضية، لكنه قد يواجه صعوبة مع المسائل الأكثر تعقيداً التي تتجاوز مستوى مسابقات الرياضيات الأمريكية AMC المستوى 12، خاصة في الهندسة. في الختام، يُعتبر إصدار NuminaMath 7B TIR أداة قيمة لمن يشاركون في التحديات الرياضية عالية المستوى، ويبرز إمكانيات الذكاء الاصطناعي في إحداث ثورة في حل المسائل الرياضية.

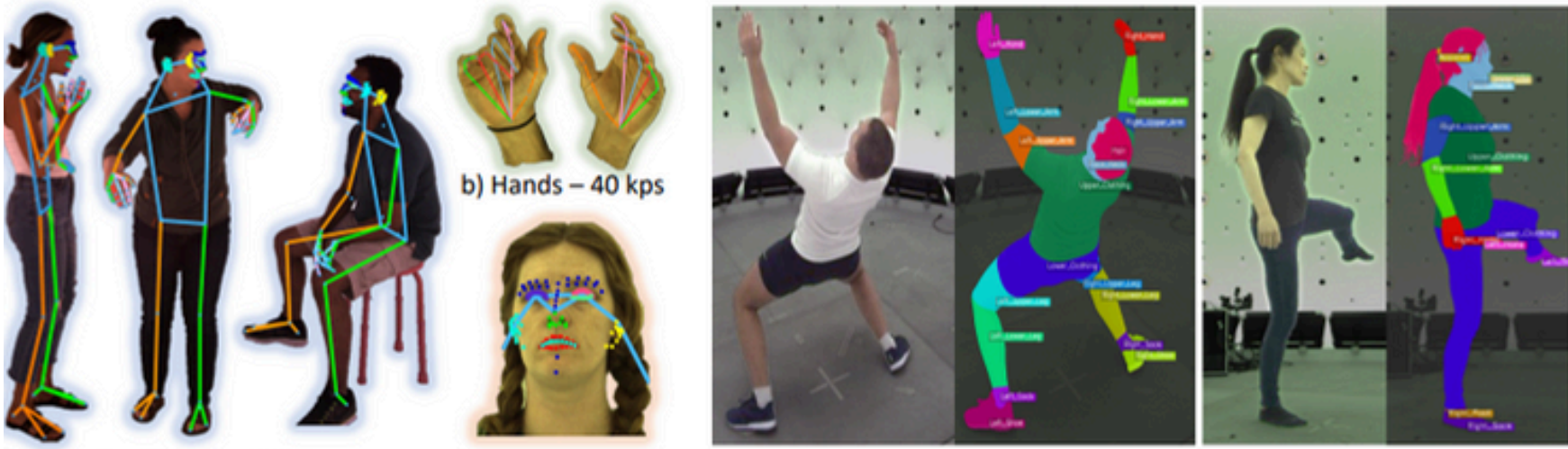
إعادة تعريف الذكاء الاصطناعي الموجه نحو الإنسان لفهم بصري من الجيل القادم

Model	#Params	FLOPs	Hidden size	Layers	Heads	Batch size
Sapiens-0.3B	0.336 B	1.242 T	1024	24	16	98,304
Sapiens-0.6B	0.664 B	2.583 T	1280	32	16	65,536
Sapiens-1B	1.169 B	4.647 T	1536	40	24	40,960
Sapiens-2B	2.163 B	8.709 T	1920	48	32	20,480

مواصفات مشفر Sapiens للتدريب المسبق على مجموعة بيانات Human-300M

تمثل Vision Transformer، التي تم تطويرها بواسطة Meta's Reality Labs، تقدماً غير مسبوق في نماذج الرؤية الموجهة نحو الإنسان. تم تصميمها خصيصاً للتفوق في أربع مهام حيوية تتعلق برؤية الإنسان: تقدير الوضع، ثنائي الأبعاد، تقسيم أجزاء الجسم، تقدير العمق، وتوقع الأسطح العادية.

- تتميز Sapiens بقدرات استدلال عالية الدقة (1K) وقابلية للتوسع، مع نماذج تتراوح بين 0.3 إلى 2 مليار من المعاملات، متفوقة بشكل كبير على النماذج الحالية غير مختلف المعايير المتعلقة بالإنسان.
- باستخدام مجموعة بيانات منسقة تحتوي على 300 مليون صورة بشرية في البيئات الطبيعية، تُظهر Sapiens قدرة عالية على التعميم، حتى في السيناريوهات التي تحتوي على بيانات معنونة محدودة أو مجموعات بيانات اصطناعية بالكامل. هذا يجعلها الخيار الأمثل للتطبيقات التي تتطلب تمثيلاً وفهماً عالي الدقة للبشر.
- يُظهر أداء النموذج في مهام مثل تقدير الوضع، تقسيم أجزاء الجسم، تقدير العمق، وتوقع الأسطح العادية ففزة كبيرة، مع تحسينات تصل إلى 53.5% في الخطأ الزاوي النسبي على المعايير مثل THuman2. كما تتفوق Sapiens في السيناريوهات الواقعية، حيث تتعامل مع الصور البشرية المعقدة والمتنوعة بدقة غير مسبوقة.



Ground-truth annotations for 2D pose estimation and body-part segmentation.

References:

- <https://arxiv.org/pdf/2408.12569>

