



RASM 2018

مسابقة التعرف على المخطوطات العلمية العربية التاريخية

"RASM 2018" هي مسابقة أطلقتها المكتبة البريطانية بالتعاون مع مختبر أبحاث PRIMA ومعهد آلان تورينج. أقيمت هذه المسابقة في سياق المؤتمر الدولي السادس عشر للتعرف على خط اليد (ICFHR2018).

تحتوي المكتبة على مجموعة كبيرة من المخطوطات العربية ، تضم ما يقرب من ١٥٠٠٠ عمل. تم ترقيم عدة مئات من المخطوطات كجزء من الشراكة بين المكتبة البريطانية ومؤسسة قطر، وإتاحتها في مكتبة قطر الرقمية.

تقام المسابقة على أمل إيجاد الحل الأمثل للنسخ الدقيق والآلي للأرشيف الرقمي الواسع والمتزايد، للمخطوطات العربية العلمية التاريخية المكتوبة بخط اليد داخل مكتبة قطر الرقمية. تهدف المسابقة الى تحسين إمكانية الوصول إلى هذا المحتوى الغني من خلال تمكين البحث عن النص الكامل واكتشافه، بالإضافة إلى تمكين تحليل النص على نطاق واسع.

حازت RDI على المركز الأول

تم تقسيم المسابقة إلى ثلاثة تحديات مختلفة: تجزئة الصفحة، واكتشاف خط النص ، والتعرف الضوئي على الحروف (OCR) .

تنافست RDI مع الأنظمة الراسخة المستخدمة في الصناعة والأوساط الأكاديمية:

١. OCR from RDI

٢. Tesseract (Google): 3.04, 4.0

٣. و ABBYY FineReader Engine 11 (FRE11)

٤. و Google Cloud Vision API

٥. و BenGurion University (KFCN).

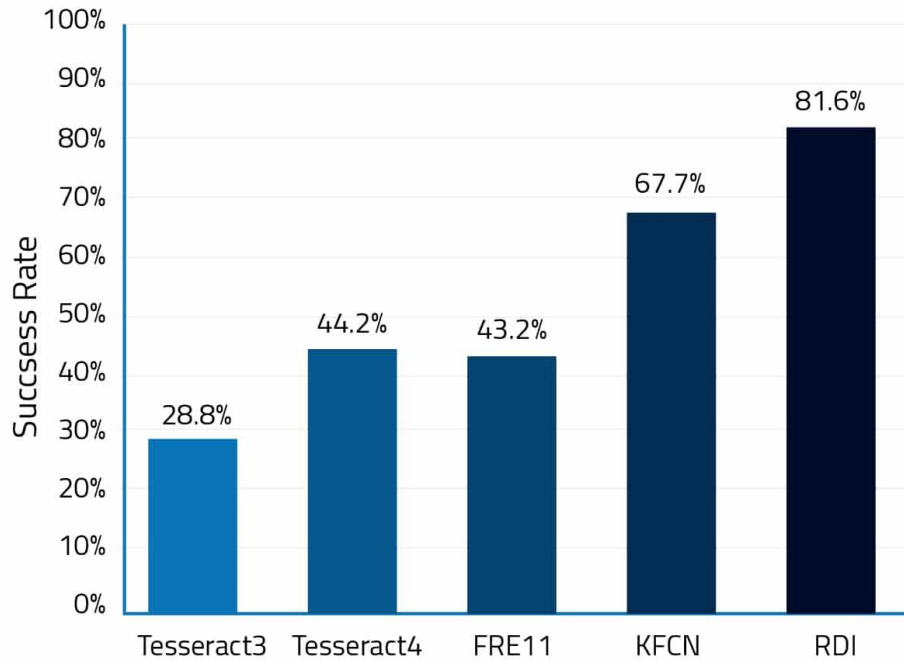
نظرًا للطبيعة الصعبة للغاية للوثائق، نعتقد أن أدائنا كان رائعاً، حيث وصلنا إلى دقة ٨١.٦٪ في "التحدي ٢ –

تجزئة سطر النص" بفارق يزيد عن ١٣٪ عن المركز الثاني. كما حققنا دقة بلغت ٨٥.٤٤٪ في "التحدي ٣ –

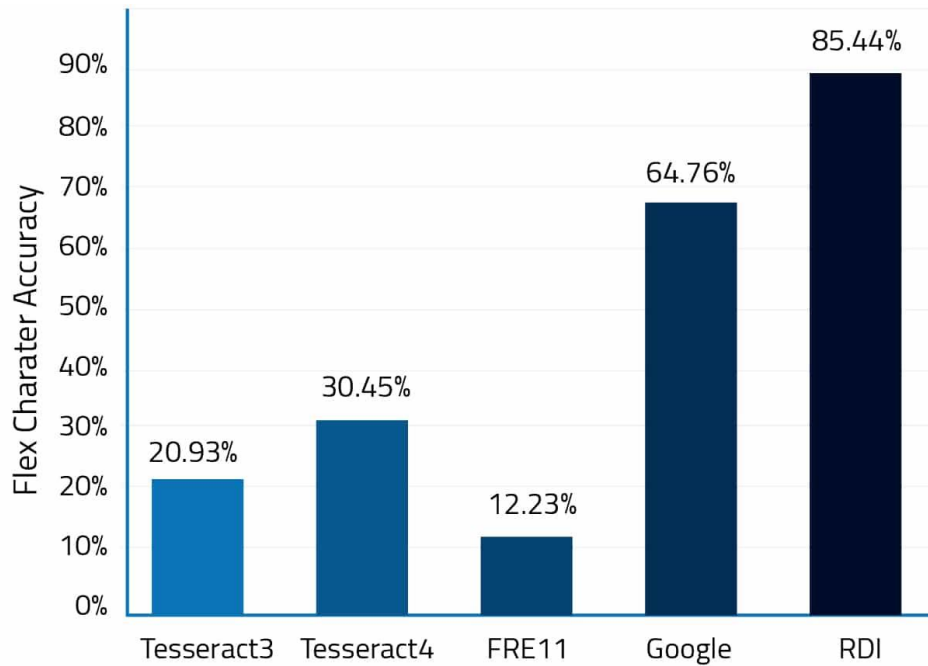
التعرف على النص" بفارق يزيد عن ٢٠٪ عن المركز الثاني.

رابط المسابقة: <https://www.primaresearch.org/RASM2018/results>

رابط ورقة البحث: https://www.primaresearch.org/www/assets/papers/ICFHR2018_Clausner_RASM.pdf



التحدي ٢ - تجزئة الخطوط من النص



التحدي الثالث - التعرف على النص