نظام متعدد الوكلاء للتعرف على المخطوطات العربية

يوسف بوليد، عبد الغني سهار، محمد يوسفي الكتاني

ملخص

في هذه الورقة سنقدم بنية للتعرف الآلي على النصوص المكتوبة بخط اليد، و سنسلط الضوء على الإشكالية الموجودة في غالبية النظم الموجودة، والمتمثلة أساسا في انتشار الأخطاء الناتجة عن التنفيذ المتتابع والمتتالي لمراحل عملية التعرف على الحروف. سوف نوضح كيف يمكن حل هذه المشكلة من خلال توفير بنية على أساس النظم المتعددة الوكلاء، الشيء الذي سيسمح بالتعاون في ما بين مختلف هذه المراحل على مختلف المستويات.

كلمات المفتاح : التعرف على الكتابة اليدوية، بنية متعددة الوكلاء، الذكاء الاصطناعي

I. الخلفية

في الوقت الحاضر، حتى مع الاستخدام المكثف للتكنولوجيات الجديدة، نجد أن الإنسان مازال يستخدم الورق كوسيلة للاتصال وتخزين المعلومات. جمع وأرشفة هذه الوثائق هي واحدة من الأهداف الكبرى للدول، حيث أنها تعتبر منجم من المعلومات، وتعد رافعة اجتماعية واقتصادية وثقافية تشهد على تطور الأمم.

يتم تخزين كميات ضخمة من الوثائق الأصلية التي لا تزال في حالة يمكن استغلالها لإنتاج نصها الإلكتروني، والتي ستمكن للمستخدمين من البحث في هذه الوثائق وذلك باستعمال كلمات نصيه عوض البحث يدويا في الصور الممسوحة ضوئيا (Photo Scanner)

النظم التي تقوم بتحويل الوثائق الورقية إلى نصوص إلكترونية تسمى نضم التعرف على الحروف المكتوبة بخط اليد (Handwritten Character Recognition) أو نضم التعرف الضوئي على الحروف (Optical Character Recognition) وفقا لنوع الوثيقة، إذا كان النص مكتوب بخط اليد أو مطبوع.

هناك العديد من التطبيقات لهذه النظم، مثل: رقمنة الوثائق، التخزين، استرجاع وفهرسة الكتب، الفرز التلقائي للرسائل، التعرف على الشيكات المصرفية و النماذج .

أهمية هذه التطبيقات أدت لبحوث مكثفة لعدة سنوات.

غالبا ما يتكون نضام التعرف على الحروف المكتوبة بخط اليد من خمس مراحل:

١. مرحلة ما قبل المعالجة(Preprocessing) : التي تتكلف بتهيئة الوثيقة للمراحل المقبلة، كإزالة الشوائب)الأوساخ و البقع(، تقويم ميل الحروف والكلمات، تصحيح انحناء السطور، تحويل الوثيقة إلى الأبيض والأسود

٢. مرحلة التقسيم (Segmentation): تقسيم الوثيقة إلى اسطر، كلمات تم حروف

٣. مرحلة تمثيل البيانات(Feature extraction) : استخراج معلومات تمكن من تمييز الحروف فيما بينها

٤. مرحلة صنع القرار (Decision-making) :التعرف على الحروف و الكلمات

٥. مرحلة ما بعد المعالجة(Post processing) : التأكد من صحة الحروف والكلمات و الجمل من خلال السياق و البحث في المعجم.

كل هذه المراحل لها عدد من الأدوات والأساليب من أجل حل مشكلة معينة، مثل: تقنيات تصفية الصور، إزالة الشوائب، تصحيح الانحراف، تقسيم النص إلى سطور وكلمات وحروف، تقنيات للتحقق من المفردات والجمل، تقنيات التعلم والتصنيف مثل شعاع الدعم الآلي (Support Vector Machines)، شبكة عصبونية اصطناعية (Neural Networks)، نموذج مارف المخفي (Hidden Markov

Models).

يهدف هذا البحث إلى محاولة إظهار وتوضيح العيب الرئيسي الموجود في معظم نظم التعرف على النصوص المكتوبة بخط اليد، وثانيا لتقديم بعض الملاحظات المستخلصة من تحليل سلوك الإنسان أثناء عملية القراءة

بناء على هذه الحقائق سنبين كيفية تصميم بنية تسمح بتقليد قدرة الإنسان على القراءة، من أجل دمجها في نظام التعرف على النص المكتوب بخط اليد، الشيء الذي سيمكن من الحصول على أفضل معدل لتحويل صورة الوثيقة لنص إلكتروني.

II. الإشكالية

معظم النظم توظف هذه الأدوات و الأساليب بطريقة خطية و متتالية. مثال لبنية خطية في الشكل ١.

الشكل ١. مراحل عملية التعرف على الحروف.

وجدنا ثلاثة نقط تمثل العائق الرئيسي:

١. من جانب التفعيل المتتالي للمراحل

يتم تنفيذ المراحل بالتسلسل، وهذا يعني أنه إذا لم يتم تنفيذ خطوة بشكل صحيح، فإنها قد تتسبب بتراكم الأخطاء بين المراحل، وذلك، يؤثر على نتيجة التعرف النهائية. كما هو مبين في الشكل أدناه، إذا فشلت عملية تحويل الصورة إلى أبيض وأسود من استرداد النص داخل منطقة نزيف الحبر (الجزء الأزرق)، فإن مرحلة إزالة الشوائب ستقوم بإزالة هذه المنطقة السوداء التي تمثل معلومات هامة للمراحل المقبلة (الشكل ٢)

الشكل ٢. متال تغيير صورة من صيغتها الأصلية الى صيغة اللون الأبيض والأسود

٢. من ناحية التعامل مع الوثيقة بالكامل

بعض الخوارزميات تحاول معالجة الوثيقة محليا، على سبيل المثال تحويل الصورة إلى أبيض وأسود باستعمال مستوى التدرج الرمادي المنبثق من تحليل الصورة في عدة أجزاء. ولكن معظم التقنيات والأساليب تطبق نفس العلاج على الوثيقة بأكملها. وهذا يمكن أن يصلح للوثيقة النظيفة التي تحتوي على أجزاء متجانسة. خلاف ذلك، فإنها ليست مناسبة للوثائق المتدهورة، حيث يمكن أن تتكون من أجزاء لا تتقاسم نفس الخصائص.

على سبيل المثال، إذا كانت وثيقة تحتوي في بعض أجزاءها على أوساخ على شكل نقط بيضاء وسوداء)تحدث أثناء عملية المسح الضوئي(، ثم إذا كانت خوارزمية إزالة الأوساخ تقدر حجم هذه الأخيرة من الوثيقة بأكملها، فسيتم احتساب شكل و نقط الحروف كأوساخ، وبالتالي سوف تمحى.

من وجهة النظر هذه، يمكننا القول أن كل جزء من الوثيقة له خصائص معينة، وبالتالي يمكن أن يعامل بالطريقة المناسبة

٣. من ناحية صيغة الألوان

توجد أساليب معينة للصور الملونة والصور الرمادية والصور بالأبيض والأسود، ولكن مرة أخرى، هذه الخوارزميات تعالج الوثيقة بأكملها، ولكن وفقا لحالتها ، قد تكون هناك مناطق غير متجانسة داخل نفس الوثيقة. وهو الأمر الذي يعني أنه من أفضل لمنطقة معينة، استعمال الصيغة المناسبة التي تسمح باستخراج المعلومات المفيدة، سواء في صيغة الرمادي أو صيغة الألوان أو صيغة الأبيض والأسود.

III. عملية قراءة الإنسان

عندما يقوم شخص ما بقراءة وثيقة مكتوبة بخط اليد لأول مرة، فإنه يوظف عدة مستويات من التحليل بطريقة سريعة جدا حتى دون إدراك لذلك. من خلال القيام بهذا، يعتقد انه يتعرف على معنى الكلمات على أساس التعرف على شكل الأحرف

ستانيسلاس ديهان [١] يدعي أن في حالة القراءة، يتم تشغيل ثلاث مجموعات من وظائف المخ: شبكات التعرف البصري الثابتة المعنية بالتعرف على الكتابة يليها تحويل الحروف إلى الصوت والوصول لمعنى الكلمات والجمل.

كما ورد في [٢] لدينا العديد من الاستراتيجيات لقراءة وثيقة، مثل قراءة كاملة (قراءة كلمة بعد كلمة)، التفتيش (البحث في منطقة محددة من الوثيقة)، وأخيرا النظرة العامة (معاينة الوثيقة)

في الواقع يمكننا تحديد، من بين أمور أخرى، بعض من الآليات التي تدخل في عملية القراءة، مثل:

- استخراج بعض المعلومات من الوثيقة، مثل الخصائص الهندسية للنص، أسلوب الكتابة، ونوعية الورق ... الخ.
- كشف المناطق المتسخة من المناطق النظيفة.
- تجاهل الأوساخ أو جزء من كلمة أو حرف.
- اقتراح كلمة أو حرف مكتوب بشكل سيئ .
- استبدال حرف مفقود.
- اقتراح تقسيمات لحروف كلمة مبهمة
- البحث في القاموس في صحة عبارة ما.
- استخراج المفهوم الدلالي (semantic) للكلمة من سياقها (الجملة)
- تأخير قرار بشأن شكل كلمة مبهمة، حتى يتم جمع مزيد من المعلومات أثناء القراءة من أجزاء أخرى من الوثيقة.
- مقارنة بين الكلمات والحروف معا لإزالة الغموض في بعض الحالات
- تعلم أسلوب الكتابة للكاتب (الطريقة التي تتم فيها كتابة الحروف وفقا لموقعها) من خلال قراءة أجزاء مختلفة من النص

من هذه الملاحظات، يمكننا القول أن هناك مجموعة من استراتيجيات القراءة التي تنطوي على:

التوازي:

الوصول إلى مستويات سياقية (contextual level) مختلفة في الآن نفسه: الرسم (graphic)، الشكل (morphological)، النحو (syntactic)، الدلالة (semantic) والمعجم (lexical)

التعاون:

الوصول إلى مستوى سياقي محدد، على سبيل المثال، عندما يكون هناك خطأ في الدلالة، يمكن أن نعود إلى مستوى أدنى لمعرفة أين

تكمن المشكلة. كما أنه يمكن استنتاج المعلومات من بعض المناطق لاستخلاص فرضيات من شأنها أن تساعد في اتخاذ القرارات في مواضع أخرى (التكاملية بين المناطق)

الصيغة اللونية:

في بعض الأحيان عندما نواجه كلمة غير مفهومة، يمكننا جلب الورقة بالقرب من أعيننا (كما لو استعملنا مكبرة، لمعرفة المزيد من التفاصيل). أو لوضع الورقة بعيدا عن أعيننا من أجل أن نرى كلمة في سياقها. يمكن أن نعتبر هذا بمثابة التحجيم (scaling)

كل هذه الآليات والاستراتيجيات المتبعة في هذه العملية تعطي الإنسان القدرة على قراءة وثائق لم يرها من قبل.

الملاحظات الأولى والحقائق حول عملية القراءة تحتنا إلى إعادة النظر في أهمية بنية نظام التعرف على الحروف المكتوبة بخط اليد نحتاج إلى تقسيم الوثيقة بطريقة ذكية، للتعامل مع كل منطقة بالطريقة الأنسب، من حيث صيغة اللون والعلاج المناسب وتوفير لكل مرحلة ذاكرة لتتبع العلاجات والسماح بالعودة خطوة إلى الوراء لاستخراج معلومات أخرى وتصحيح الأخطاء

من وجهة النظر هذه، يمكننا القول أن أفضل نظام للتعرف على الكتابة اليدوية هو الذي يقلد عملية قراءة الإنسان.

وجدنا أن الأنظمة امتعددة الوكلاء (Multi Agents Systems) هي الأكثر ملائمة لهذا النوع من المشاكل

للوكلاء القدرة على التواصل مع وكلاء آخرين في نفس البيئة (ونقصد بهذا الصورة)، ولديهم سلوك ذاتي يتيح لهم التصرف وفقا لهدف خاص بهم ومعرفة البيئة المحيطة بهم

IV. الأعمال ذات صلة بالموضوع

في [٢] تم اقتراح نظام متعدد الوكلاء للتعرف على نص خط اليد، وذلك لتكييف التعرف على خط يد الكاتب.

النص مجزأ إلى أجزاء الحروف، يتميز كل واحد منها بميزات شكلية و رمزية. للتعرف على حرف يتم تفعيل التعرف على مستوى الكلمة، ومن ثم يمكن أن تستمد المعلومات السياقية لكل حرف و أيضا من خلال الحروف المتشابهة الشكل. هذه المعلومات السياقية يمكن أن توفر فرضيات رمزية لميزات ثابتة تميز الكاتب.

النظام مركب باستخدام بنية مفتوحة تسمى EMAC، التي توفر قدرات الاتصال والسيطرة وضمان تقاسم مساحة عمل مشتركة. الوكلاء هم مجموعات من الخبراء في التحليل، تجزئة النص، والتعرف على أجزاء الحروف و الكلمات.

ومع ذلك، تم افتراض أن مرحلة التقسيم و مرحلة ما قبل المعالجة تمت بشكل صحيح، و أن النظام يبدأ بعد هذه الخطوات. ولكن هذا لا يمنع من تسبب أخطاء إذا كانت حالة الوثيقة متدهورة.

V. النهج المقترح

الهدف هنا هو محاكاة عملية قراءة الإنسان عن طريق استخراج الخصائص والآليات التي تشارك في هذه العملية.

هناك حاجة إلى بنية تسمح لهيكلة ودمج الأساليب والأدوات بطريقة متدرجة وتمكن من ضمان التعاون بين المراحل المختلفة لعملية التعرف على الكلمات المكتوبة بخط اليد.

لنكون قادرين على محاكاة عملية قراءة الإنسان، نقترح بنية تقوم على نظام متعددة الوكلاء وذلك لتنفيذ خوارزميات على أساس الفرضيات المذكورة أعلاه بسهولة.

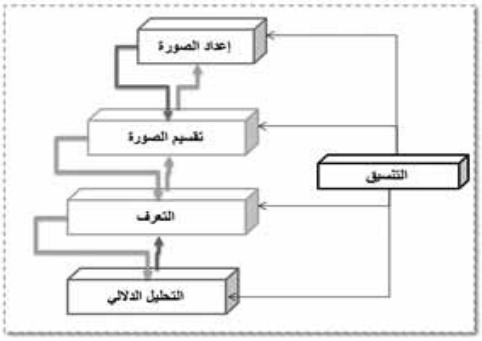
البيئة ثابتة و يمكن ملاحظتها بشكل كامل ولديها ثلاث صيغ: باللون، بالمستوى الرمادي، بالأبيض وأسود (الشكل ٣)

الشكل ٣. الصيغ المختلفة للوثيقة

وتنقسم الصورة إلى أجزاء، وداخل كل جزء لدينا مجموعة من الوكلاء (كما هو مبين في الشكل ٤:(

- وكيل مسئول عن إعداد الصورة (Normalization) ، الذي يحتوي على وظائف لتنقية الصورة وتهيئ وتصحيح النص.
- وكيل مسئول عن تقسيم الصورة (Segmentation) إلى أسطر وكلمات وأحرف.
- وكيل مسئول عن التعرف(Recognition) ، الذي تم تدريبه على تعرف الأحرف والكلمات.
- وكيل مسئول عن التحليل الدلالي (Semantic) ، والذي يحتوي على وظيفة التحقق من الكلمات في القاموس، وتصحيح المفردات.
- وكيل مسئول على الأداء الكلي للوكلاء أعلاه (Control) وعن التنسيق فيما بينهم

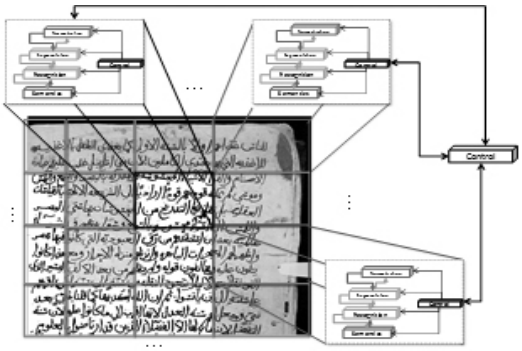
يحتوي كل وكيل على مجموعة من الوظائف مع ذاكرة تمكنه من الرجوع عن بعض الإجراءات التي قام بها وفقا لمقياس التقييم. على سبيل المثال، الوكيل المسئول عن إزالة الشوائب، يمكن أن يحتفظ على أجزاء تمت إزالتها في حالة أن وكيل آخر أراد معرفة ما اذا كانت مفيدة لنتيجة التعرف.



الشكل ٤. تنضيم الوكلاء

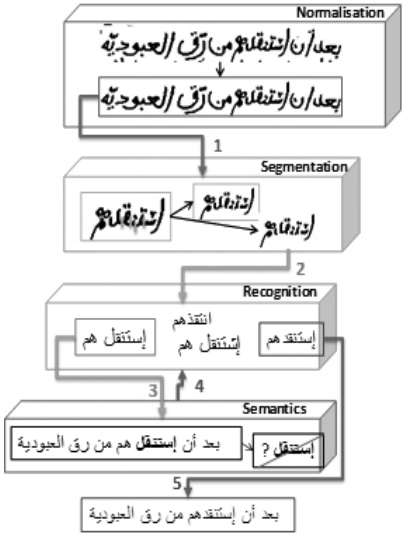
قد لا تكون لأجزاء الوثيقة نفس الأبعاد، وفقا لتجانسها، فإنه يمكن تقسيمها إلى أجزاء فرعية يسيطر عليها وكلاء فرعين. كل جزء له طبيعة محددة لذلك يمكن أن يعامل بطريقة أو بأخرى. وهذا ما يمكن من العمل على نفس الوثيقة على صيغ مختلفة في مناطق مختلفة

يتم التعاون بين الوكلاء و تبادل المعلومات من مختلف الأجزاء بطريقة مركزية من خلال وكيل التنسيق المركزي الذي يتوفر على رؤية كلية لصورة الوثيقة (الشكل ٥)



الشكل ٥. البنية المقترحة، والتي تبين التعاون بين الوكلاء في مناطق مختلفة من الوثيقة

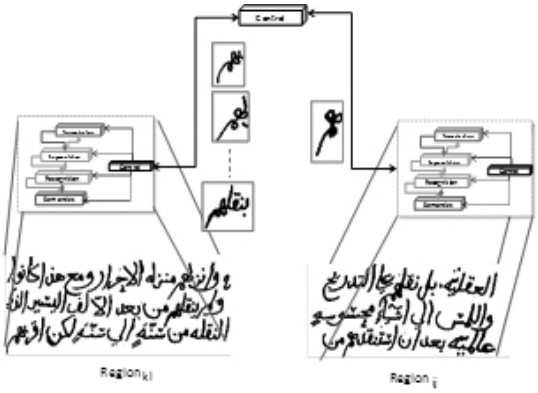
يعمل الوكلاء بشكل مستقل، ويمكن أن يتعاونون فيما بينهم ضمن الجزء ذاته عند مستويات سياقية، كما هو الحال في حالة حرف مبهم، يمكن لوكيل التعرف أن يطلب لوكيل التقسيم منحه المزيد من الفرضيات حول نقاط تقسيم أخرى، أو أن يطلب من وكيل التنقية اذا كان قد أزال شيء يمكن أن يكون جزء من حرف. وحالة أخرى هي أنه يمكن للوكيل الدلالي أن يطلب من وكيل التعرف مزيد من الكلمات الموجودة في القاموس لاختيار الكلمة الصحيحة التي تتماشى مع سياق الجملة (الشكل ٦)



الشكل ٦. مثال للتعاون بين وكلاء الدلالة والتعرف

التعاون يمكن أن يكون أيضا بين وكلاء من نفس النوع في أجزاء مختلفة، على سبيل المثال يمكن لوكيل التعرف طلب من وكلاء آخرين نتيجة التعرف على حرف معين في سياق آخر، أو أن يؤخر قراره حتى تجتمع مزيد من المعلومات. أيضا من أجل التمييز بين الشوائب و الحروف، وكيل التنقية يمكن أن يبحث في الطريقة التي يتم فيها كتابة حرف ما في أجزاء أخرى (الشكل ٧(

ويمكن اعتبار هذا التعاون بين الوكلاء كتعاون بين أجزاء مختلفة من الوثيقة الشيء الذي يعطي الوكلاء القدرة على تعلم الوثيقة نفسها وتكيف قرارات التعرف وفقا لنمط الكاتب.



الشكل ٧. التعاون بين وكلاء من مختلف المناطق

البنية المقترحة في هذا البحث تسمح بدمج جميع مراحل التعرف على الأحرف بطريقة تعاونية وذلك للتغلب على المشاكل المتعلقة بفقدان المعلومات في النهج المتسلسل

VI. الحالة الراهنة للنظام

بدأنا برمجة وكيل التنقية [٤]. نظرا لتنوع أنواع الشوائب الموجدة في المخطوطات التاريخية، وجدنا أنه من المناسب استخراج محتوى النص مع تجاهل الشوائب ولاحظنا أن هدا الذي يفعله الإنسان في عملية القراءة (القارئ لا يبحث عن أوساخ محددة لإزالتها. بل يتجاهل نوع وطبيعة الشوائب ويبحث عن مضمون النص في الصورة وضمن الأوساخ(

الخط العربي متصل ويمكن للحروف أن تكون في شكل خطوط ومنحنيات وحلقات، ويمكن لهذه الحروف أن تلتصق معا بعضها أفقيا. غالبا ما تمثل نقط الاتصال هذه تقاطعات.

ووفقا لهذه المعلومات، وجدنا أن مفهوم التقاطعات في مناطق التصاق الأحرف تسمح بالتمييز بين مكونات النص) حرف أو جزء من حرف(والشوائب. في الواقع، كلمات النص تحتوي على كمية أقل من التقاطعات

لحساب عدد التقاطعات في النص، اقترحنا طريقة تسمح لنا بتصنيف العناصر إلى: نص، نقط أو أوساخ

وهكذا، فإن خوارزمية الكشف عن النص وإزالة الشوائب، تبدأ بإزالة النقط البيضاء والسوداء بناء على متوسط عرض الأحرف، ثم تصنف المكونات من أجل إزالة الأوساخ وأخيرا، يتم إدراج النقط إلى أحرفها باستخدام موقعها وقربها من هذه الأخيرة

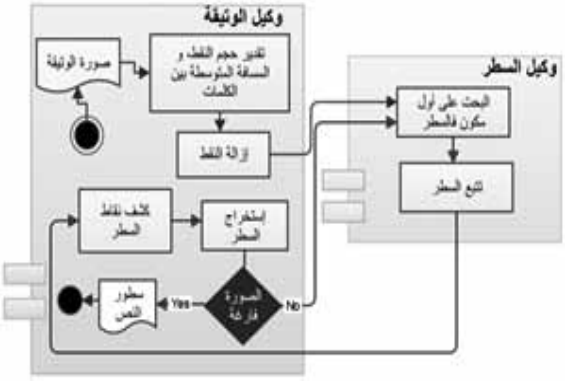
نتيجة هذه العملية مبينة في الشكل ٨.

الشكل ٨. نتيجة إستخراج النص وإزالة الشوائب

قمنا مؤخراً بإرسال مقال عن استخراج سطور النصوص العربية المكتوبة بخط اليد، استنادا إلى آليات وتقنيات عملية القراءة البشرية للكشف عن السطور. اقترحنا نهج قائم على نضام متعدد الوكلاء لكشف وتجميع المكونات المتصلة التي تنتمي إلى نفس السطر من خلال الاستفادة من ميزات وترتيب تلك المكونات.

والفكرة هي إعادة بناء سطور الوثيقة، وذلك بكشف لكل سطر، مكوناته المختلفة (حرف، جزء من الكلمة أو كلمة كاملة) بطريقة متتالية. لهذا أنشأنا وكيلين أولهما يدعى وكيل الوثيقة وهو المسئول عن استخراج السطور والثاني يدعا وكيل السطر، وهو المسئول عن الكشف عن مكونات السطر.

لاستخراج السطر الأول، الوكيل يبدأ الكشف عن المكون الأول في السطر، تم ينتقل إليه، ثم يكتشف المكون التالي ،تم ينتقل إليه وهكذا دواليك، حتى يصل إلى نهاية الخط. بعد ذلك يتم استخراج وإزالة تلك المكونات من الصورة واعتبار السطر الثاني كالسطر الأول، ويتم تنفيذ هذه العملية بشكل متكرر حتى يتم استخراج كل السطور . الشكل ٩ يبين رسم تخطيطي يوضح مهام وصلات هذين الوكيلين



الشكل ٩. رسم مبياني يوضح العلاقة بين الوكيلين ومهامهم

ويوضح الشكل ١٠ بعض نتائج النظام المقترح لاستخراج سطور نصوص اللغة العربية المكتوبة بخط اليد

الشكل ١٠. بعض النتائج إستخراج سطور النص

VII. الخلاصة

في هذا العمل، اقترحنا بنية للتعرف على النص العربي المكتوب بخط اليد على أساس مفهوم النظم المتعددة الوكلاء، والذي يسمح لإعطاء مزيدا من المرونة في تنفيذ استراتيجيات القراءة البشرية، وبالتالي يمكن أن يزيد من فرص تحسين معدل التعرف.

في المستقبل سوف نكمل تنمية الوكلاء الآخرين باقتراح نظم جديدة أو تكييف النظم الموجودة بحيث يمكن دمجها بسهولة في البنية المقترحة.

VIII. المراجع

[١] Dehaene, S. (٢٠٠٧). Neurones de la lecture (Les): La nouvelle science de la lecture et de son apprentissage. Odile Jacob.

[٢] Eglin, V., Bres, S., & Emptoz, H. (١٩٩٩). Structuration de documents par repérage de zones dintérêt. TS. Traitement du signal, ١٦(٣), ٢١٧-٢٣٩.

[٣] Heutte, L., Nosary, A., & Paquet, T. (٢٠٠٤). A multiple agent architecture for handwritten text recognition. Pattern recognition, ٣٧(٤), ٦٦٥-٦٧٤.

[٤] Boulid, Y., & Elkettani, M. Y. (٢٠١٤). Approach for Arabic Handwritten Image Processing: Case of Text Detection in Degraded Documents. International Journal of Computer Applications, ١٠١(١٤).