دراسة تطبيق خوارزمية دوغلاس بيوكر في تبسيط النص المكتوب بخط اليد

مصطفى علي ابوزريدة[1]، سالمة إبراهيم الفتيوري[2]، أسماء محمد السمو[2]، أمل سليمان الفطيسي[2]، رقية رمضان أبوعجيلة[2]
قسم علوم الحاسوب، كلية تقنية المعلومات، جامعة مصراتة، ليبيا
[2] قسم الحاسوب، المعهد العالي للمهن الهندسية مصراتة، ليبيا
[1]abuzaraida@it.misuratau.edu.ly

الخلاصة: يعتبر التعرف الضوئي للنصوص من أحد فروع التعرف الضوئي باستخدام الحاسوب. حيث يلقى هذا المجال إقبالا كبيرا من البحاث حول العالم لما له من أهمية كبيرة في التعامل في إدخال النصوص بطريقة جديدة تختلف على الطريقة التقليدية وهي باستخدام لوحات المفاتيح. يمر نظام التعرف الضوئي على النصوص بمراحل مختلفة تختلف في الوظيفة والهدف، حيث تتسلسل المراحل من إدخال النص إلى المرحلة التمهيدية (Preprocessing Phase) ثم مرحلة استخلاص المعرفات (Features Extraction Phase) وأخيرا مرحلة التعرف (Recognition Phase) . تتكون المرحلة التمهيدية من بعض المراحل التي من أهمها تقليل النقاط للوصول إلى أفضل تمثيل للنقاط المكونة للنص. من أشهر الخوارزميات التي تستخدم في تنفيذ هذا الإجراء هي خوارزمية دوغلاس بيوكر Douglas-Peucker. حيث تهدف هذه الورقة إلى إجراء دراسة حول تطبيق خوارزمية دوغلاس بيوكر في تبسيط النص المكتوب بخط اليد. وتمت دراسة الفروقات في هذه الخوارزمية من خلال تحديد ثلاثة معايير مختلفة. تم تحديد الفروقات بالنسبة للإنسان من حيث الشكل العام باستخدام استبيان مصمم لهذا الغرض. أما من حيث كفاءة النظام، فيعتبر تقليص عدد النقاط من أهم العوامل المساعدة في تسريع عملية المعالجة خلال المراحل القادمة في نظام التعرف الضوئي المباشر. سيتم في هذه الدراسة كذلك استعراض شامل لمراحل التعرف الضوئي المباشر وغير المباشر على النصوص المكتوبة بخط اليد أو المطبوعة مسبقا على مستندات ورقية. وكذلك فهم الفروقات بين أنظمة التعرف الضوئي المباشر وغير المباشر.

الكلمات الجوهرية: التعرف الضوئي، الارقام العربية، التعرف الضوئي المباشر، مرحلة المعالجة التمهيدية.

1. المقدمة

خلق الله سبحانه وتعالى الإنسان وأنعم عليه بالكثير من النعم، فأنعم عليه نعمة العقل التي ميزته عن سائر المخلوقات. ومنح الله للإنسان القدرة على التعبير بما يجول في هذا العقل باستخدام لسانه كلغة يفهمها من حوله من باقي البشر. مع تطور حياة الإنسان الأول وتكوين المجتمعات البشرية، وجد الإنسان نفسه غير قادر على حفظ الكلام المتداول فاهتدى للكتابة لحفظ إنتاجه الفكري وميراثه الثقافي والعلمي من الاندثار ولتتوارثه الأجيال اللاحقة. وتدرجت مراحل الكتابة في العصور الأولى من استخدام الرسوم والنقوش لتمثيل الأحداث إلى استخدام الحروف والأرقام البدائية ووصولا إلى الكتابة بالشكل الحالي وفي شكل العديد من اللغات حول العالم. تعتبر الكتابة اليدوية طريقة مباشرة للتسجيل وإيصال الأفكار، حيث تعتبر إنتاجا للعقل واليد معا في تمثيل مرئي ومسجل على الورق وشاشات اللمس حديثا. فقد أنتجت التقنية الحديثة العديد من طرق الاتصال الكتابي، واستخدمت في عدة مجالات مثل إدخال البيانات، ومعالجة النصوص و الأرقام [1]. الكتابة اليدوية على الأجهزة المزودة بشاشات قابلة للمس سيتم استخدامها لسنوات كثيرة قادمة نظرا لتزايد الإقبال على هذه الأجهزة وتداولها في السوق، لذا فهي جديرة ببيان أهميتها، وفهم كيفية عملها وعمل الدراسات عليها للوصول إلى رضا المستخدمين. وبشكل عام يمكن استخدام الكتابة اليدوية بدلا من لوحة المفاتيح لإدخال النص أو الرقم. ويمكن كذلك الكتابة باستخدام جهاز إدخال الكتابة اليدوية مثل استخدام القلم الرقمي أو عن طريق تحريك مؤشر الماوس ثم تحويل الكلمات المكتوبة بخط اليد إلى نسق النصوص أو الأرقام المكتوبة بالحاسوب ثم يدرج النص أو الرقم في المكان المرغوب فيه [2].

2. التعرف الضوئي على الكتابة

تصنف تقنية التعرف الضوئي على النصوص من ضمن عائلة التقنيات التلقائية للتعرف على النماذج، حيث تعتمد هذه التقنيات على محاكاة بعض الوظائف البشرية في القدرة على التعرف والتمييز، حيث يتم تصميمها من أجل تسهيل وإجراء عمليات التعرف بدون تدخل الإنسان [3]. التعرف الضوئي يكون على النصوص الممسوحة الكترونيا من صيغة مكتوبة بخط اليد، أو مطبوعة على الآلة الكاتبة أو نص مكتوب على شاشة قابلة للمس وتحويلها إلى نص مرمز يفهمه الحاسوب. ويستخدم على نطاق واسع باعتباره شكلا من أشكال إدخال البيانات من المصدر الأصلي، سواء أكانت وثائق، أو إيصالات مبيعات، أو بريد أو أي نوع من السجلات المطبوعة. لان حوسبة النصوص المطبوعة أصبحت مطلبا ضروريا لكي يمكن أن يكون البحث عن النصوص أو تخزينها أو عرضها أو استرجاعها إلكترونءئبزيا أكثر مرونة حيث يمكن استخدام هذه التقنية في عمليات آلية مثل الترجمة الآلية، تحويل النص إلى كلام والتنقيب في النصوص وغيرها من الاستخدام قيد الدراسة مؤخرا [4]. تنقسم أنظمة التعرف الضوئي على النصوص المكتوبة إلى نظامين أساسيين هما:التعرف المباشر Online ونظام التعرف الغير مباشر Offline [5].

أ- التعرف الضوئي المباشرOnline

يتم التعرف على الكتابة اليدوية مباشرة أثناء الكتابة باستخدام قلم ضوئي على شاشة خاصة موصولة بالحاسوب أو باستخدام الفأرة للكتابة يدويا على الشاشة مباشرة أو بالكتابة مباشرة على شاشة قابلة للمس، يمكن لهذا النظام التسجيل اللحظي لكيفية وتسلسل عملية الكتابة، مما يسهل عملية التعرف على النصوص ومعرفة الطريقة التي كتبت بها. إن تطبيقات هذا النوع ليست كثيرة إلا أنها أصبحت قيد الدراسة مؤخرا في العديد من اللغات من قبل شركات صناعة الحواسيب وأجهزة الهواتف الذكية حيث تشهد صناعة هذه الأجهزة تطورا كبيرا وتلقى رواجا كبيرا بين المستخدمين [1].

ب- التعرف الضوئي غير المباشر Offline

في هذا النوع يتم التعرف على النص المكتوب (مطبوعا أو يدويا) على الورق بعد مسحه ومعالجته وتخزينه في الحاسوب على هيئة صور. ويعتبر هذا النوع الأكثر قدما فالعديد من البحوث والدراسات أجريت على العديد من اللغات الإنسانية. ووصلت نتائج التعرف إلى مستويات مرتفعة في بعض اللغات اللاتينية وفي التعرف على الحروف والأرقام المنفصلة [4].

3. التعرف على النص العربي

من ثمانينيات القرن الماضي أصبح هناك اهتمام متزايد نحو التعرف على الأحرف العربية. حيث يعتبر مجال مهم ليس فقط للدول الناطقة بالعربية بل حتى للدول التي لها لغات مشابهة للعربية مثل لغة الأوردو واللغة الفارسية والمالاوية وغيرها من اللغات التي لها مجموعة حروف مشابهة ومشتركة مع اللغة العربية [6]. من الملاحظ أن التقنيات المستخدمة للتعرف على الحروف اللاتينية والصينية لا يمكن تطبيقها بسهولة على اللغة العربية وذلك لطبيعة الكتابة العربية. منها اتصال الحروف العربية لتكوين الكلمات وكذلك اتجاه الكتابة واختلاف شكل الحرف بناءً على موقعه في الكلمة، وغيرها من العوامل التي أثرت في تدني نتائج الدراسات حول اللغة العربية [7].

4. مراحل التعرف الضوئي المباشر

تتكون أغلب أنظمة التعرف على النصوص الضوئية المباشرة من مجموعة من المراحل المتبعة وهي:

أ. مرحلة إدخال النص:

يقوم المستخدم في هذه المرحلة بكتابة النص على جهاز إدخال موصول بالحاسوب أو عن طريق الكتابة مباشرة على شاشة الحاسوب في الحواسيب الحديثة. تتم عملية الكتابة بواسطة قلم معد لذلك او باستخدام الإصبع. ويتم تخزين هذه الكتابة تلقائيا في ملف نصي في هيئة ثنائيات من (x)،y يتم استخدامها لاحقا في المراحل المكونة للنظام [2-8].

ب. المرحلة التمهيدية Phase Preprocessing

تسمى أيضا بمرحلة ما قبل المعالجة، حيث يتم في هذه المرحلة العمل على تقليل التشوهات التي قد تصاحب عملية إدخال النص في مرحلة الإدخال. كذلك تهدف هذه المرحلة إلى الحصول على نمط موحد للنصوص المدخلة لضمان الحصول على

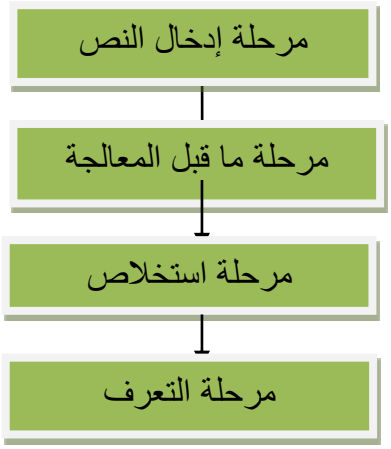
نتائج واضحة في مرحلة التعرف على النص [1]. هناك بعض الخطوات المتبعة والمكونة لهذه المرحلة منها تنعيم وصقل النص وتحجيمه وكذلك إجراء عملية إزاحة الرقم إلى نقطة الأصل.

ج. مرحلة استخلاص الخصائص Feature Extraction Phase

في هذه المرحلة يتم استخلاص بعض الخصائص التي تميز كل نص عن النصوص الأخرى. هناك أكثر من نوع للخصائص المستخدمة في أنظمة التعرف الضوئي منها ما يعتمد على الشكل العام للنص ومنها ما يستخلص باستخدام الإحصائيات حول النص المراد التعرف عليه أو من خلال هيكلية النص على سبيل المثال اتجاه حركة القلم أثناء الكتابة [5].

د. مرحلة التعرف والتصنيف Classification Phase

يتم في هذه المرحلة تعرف النظام على النص من خلال المعرفات التي تم الحصول عليها في المرحلة السابقة. يمكن استخدام خوارزميات الذكاء الاصطناعي لاستنتاج ومعرفة النص المدخل مثل الشبكات العصبية، خرائط ماركوف، طريقة شجرة اتخاذ القرار، وغيرها من تقنيات الذكاء الاصطناعي التي لا زالت قيد البحث العلمي [9]. الشكل (1) يلخص خطوات التعرف الضوئي في الانظمة المباشرة.



الشكل (1): خطوات التعرف الضوئي المباشر للنصوص [3]

5. أهداف الدراسة

تهدف هذه الورقة لإنجاز دراسة حول خوارزمية دوغلاس بيوكر التي تستخدم في تقليص وتنسيق النقاط للنصوص وكذلك تهدف إلى:

1- معرفة المعيار الأفضل عند استخدام هذه الخوارزمية عند التعرف على الأرقام المدخلة في النظام الضوئي.

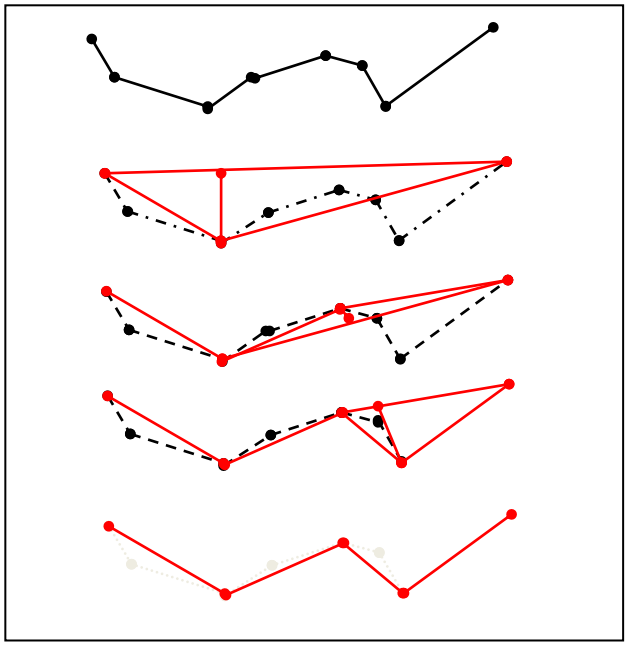
2- طريقة عمل الأنظمة الضوئية وكيفية عملها.

3- إنشاء قاعدة بيانات للأرقام العربية المجمعة من عدة أشخاص.

6. خوارزمية دوغلاس- بويكر Douglas- Peucker

هي أحدى الخوارزميات المستخدمة في تبسيط وتقليص عدد النقاط في الخطوط المستقيمة والمنحنيات. تستخدم عادة في معالجة الخرائط الجغرافية ومعالجة الرسوم. تم استخدامها أيضا في مجال التعرف على الكتابة الغير مباشرة [10]. تعمل خوارزمية دوغلاس بويكر Douglas-Peucker من خلال افتراض خط وهمي بين أول وآخر نقطة في مجموعة نقاط من نقاط المنحنى. الخوارزمية تحدد النقطة الأبعد عن هذا القطاع باعتبار النقطتين الأولى والأخيرة نقطتي نهاية. وعلى الرغم من ذلك، إذا كانت هذه النقطة أو كل النقاط الأخرى بين نقطتي البداية والنهاية هي أقرب من مسافة معينة، فإنه يزيل كل هذه النقاط بين النقطتين. أما إذا كان بعد هذه النقطة عن الخط الوهمي اكبر من قيمة معينة تعرف باسم "معيار التسامح أو

التفاوت"، فانه يتم إدراج هذه النقاط ضمن نقاط المنحنى. الشكل (2) يوضح كيفية عمل الخوارزمية على مجموعة من النقاط [10].



الشكل (2): طريقة عمل خوارزمية Douglas- Peucker [10]

لدراسة خوارزمية دوغلاس بويكر، تم تناول المرحلتين الأوليتين من نظام التعرف الضوئي المباشر على مجموعة الأرقام العربية المفردة من 0 إلى 9 وهما مرحلة إدخال النص ومرحلة ما قبل المعالجة.

7. خطوات التجربة

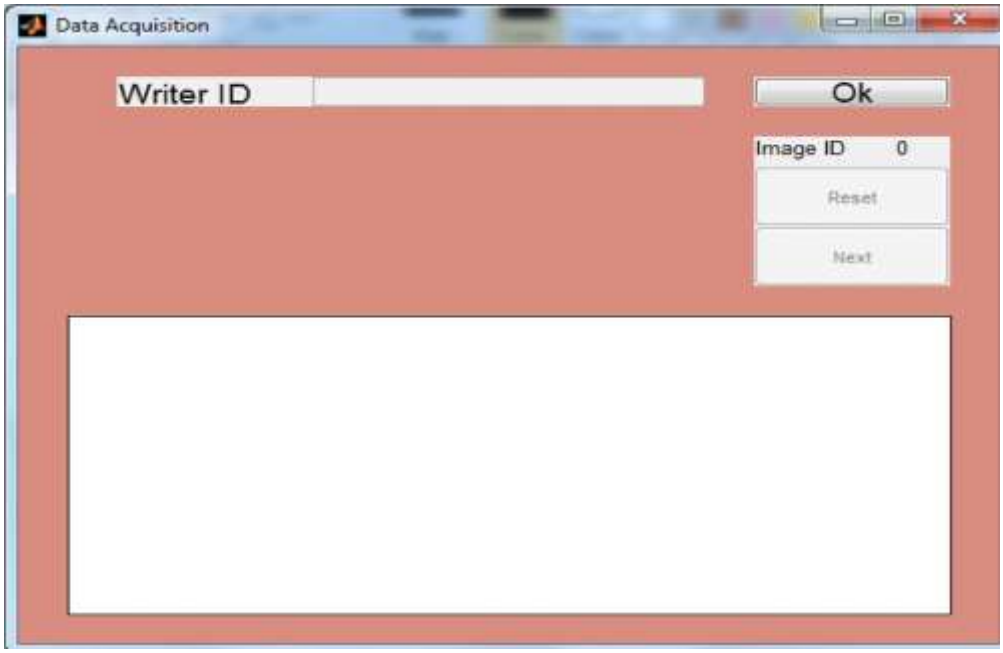
لإنجاز أهداف الدراسة بالشكل الامثل، تم اتباع الخطوات التالية:

أ. إدخال النص:

تم اختيار مجموعة من المتطوعين بشكل عشوائي للقيام بكتابة الأرقام العربية بشكل متسلسل من (0 إلى 9) عن طريق الكتابة باستخدام الاصبع او بالقلم الخاص باستخدام حاسوب محمول مزود بشاشة قابلة للمس. حيث تم استخدام حاسوب لوحي من نوع أيسر (Aecer) 1.5GHz core i3 كما هو موضح بالشكل (3).

الشكل (3): الجهاز المستخدم في الدراسة لتسجيل الكتابة بالقلم الضوئي

تتم عملية إدخال النصوص عن طريق نافذة إدخال خاصة كما هو موضح بالشكل (4).



الشكل(4): شاشة إدخال الكتابة

يتم إدخال البيانات عن طريق الخطوات التالية:

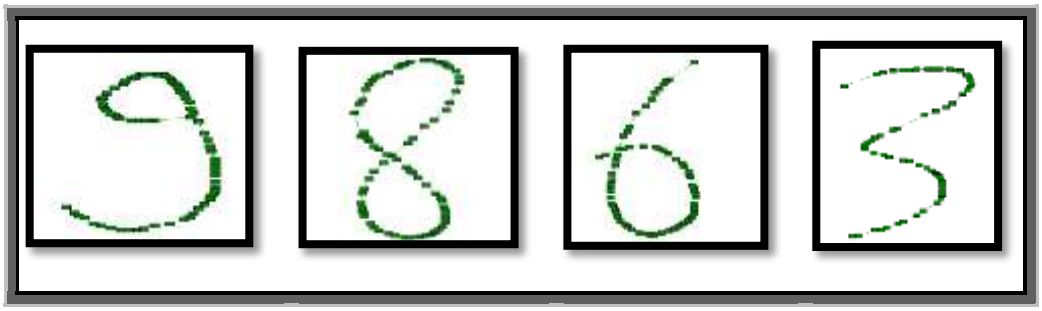
1- يجب على كل كاتب إدخال معرف الكاتب لتمييزه عن باقي الكتاب، ويتم ذلك في مربع النصWriter ID . ثم الضغط على زر موافق (OK) عند انتهاء الكاتب من كتابة المعرف.
2- يقوم الكاتب بكتابة الأرقام التي ستتم معالجتها في المربع النصي الخاص بالكتابة.
3- تتم كتابة الرقم الظاهر في أعلى يمين شاشة الإدخال Image ID.
4- عند الانتهاء من الكتابة وعدم وجود أي مشاكل في الرقم المُدخل يتم الضغط على زر التالي Next.
5- في حالة حدوث أي خطأ عند إدخال الرقم يتم الضغط على الزر Reset لإعادة كتابة الرقم مرة أخرى بشكل صحيح.
6- يتم كتابة باقي الأرقام بنفس الكيفية إلى غاية الرقم 9.

يتم تخزين الأرقام المكتوبة من قبل الكاتب في ملفات نصية كلا على حدا في مجلد واحد مسمى بمعرف الكاتب. حيث سيتم استخدامها لاحقا. الشكل (5) يبين كيفية تخزين حركة الكتابة في ملف نصي في صيغة ثنائيات من الإحداثيات x)،(y.

E	D	C	B	A	
			0.261905	0.056292	13
			0.277778	0.076159	14
			0.293651	0.09106	15
			0.317461	0.109272	16
			0.325397	0.114239	17
			0.361111	0.124172	18
			0.40873	0.135762	19
			0.420635	0.135762	20
			0.353175	0.124172	21
			0.230159	0.109272	22
			0.095238	0.100994	23
			0.079365	0.102649	24
			0	0.102649	25
					26

الشكل(5): كيفية تخزين البيانات أثناء عملية الكتابة

يمكن استعراض إحداثيات الكتابة لاستعراض النص المدخل باستخدام برنامج Microsoft excel كما هو موضح بالشكل(6).



الشكل (6): عينات من الأرقام المدخلة

ب. **مرحلة ما قبل المعالجة Preprocessing Phase:**

بعد إنجاز عملية إدخال الرقم وتخزينه على هيئة ثنائيات (x،y) في ملف نصي، يتم إجراء بعض الإجراءات والخطوات المطلوبة في المرحلة التمهيدية (ما قبل المعالجة). يمكن تلخيص الخطوات في النقاط التالية:

تحجيم الرقم المدخل:

يتم تغيير حجم الأرقام المدخلة في هذه الخطوة إلى حجم موحد. حيث تسهّل هذه الخطوة في التعرف على الأرقام والنصوص في مرحلة المعالجة. تساهم هذه الخطوة من رفع معدل التعرف على الأرقام والنصوص وذلك بإعطاء حجم موحد لكل رقم مدخل من كل كاتب.

تبسيط نقاط الرقم Digit Simplification:

يعتبر هذا الإجراء من الإجراءات المتبعة والمشكّلة لمرحلة ما قبل المعالجة. يتم تنفيذ هذا الإجراء للتقليل من عدد النقاط المكونة للرقم المدخل بحيث يتم تبسيط الرقم أو النص وسهولة التعامل معه وتقليل زمن التعرف على النص في المراحل القادمة من نظام التعرف الضوئي.

تم استخدام خوارزمية Douglas- Peucker في عملية تقليص وتبسيط الأرقام في هذه الدراسة، حيث تعتبر أحدى الخوارزميات المستخدمة في تبسيط وتقليص عدد النقاط في الخطوط المستقيمة والمنحنيات. تستخدم عادة في معالجة الخرائط الجغرافية ومعالجة الرسوم. تم استخدامها أيضا في مجال التعرف على الكتابة غير المباشر.

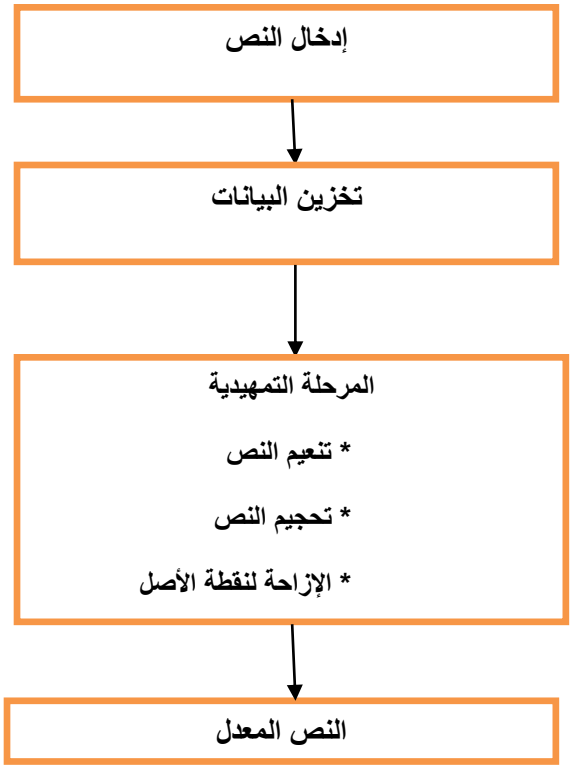
بشكل عام يمكن تلخيص خطوات خوارزمية دوغلاس Peucker Douglas في الخطوات التالية:

- قراءة نقاط نص مكتوب بخط اليد في هيئة ثنائيات (x،y).
- إنشاء شكل متعدد الخطوط بين نقطتي البداية والنهائية. وهذا هو الشكل التقريبي الأوّلي المبسط للنص.
- يتم إيجاد المسافة بين الخط المستقيم والنقاط المتبقية. فإذا كانت النقطة المختبرة تبعد أكثر من معيار التسامح والتفاوت المحدد فإنها تضاف الى نقاط التبسيط.
- يتم تكرار الخطوة 3 حتى نهاية تسلسل النقاط.

الإزاحة لنقطة الأصل:

بعد الانتهاء من إجراء الخطوات السابقة، يتم إجراء عملية إزاحة الرقم إلى نقطة الأصل. حيث تساعد هذه الخطوة في توحيد مرجعية النقاط للأرقام والنصوص.

يمكن تلخيص خطوات الدراسة كما بالشكل (7).



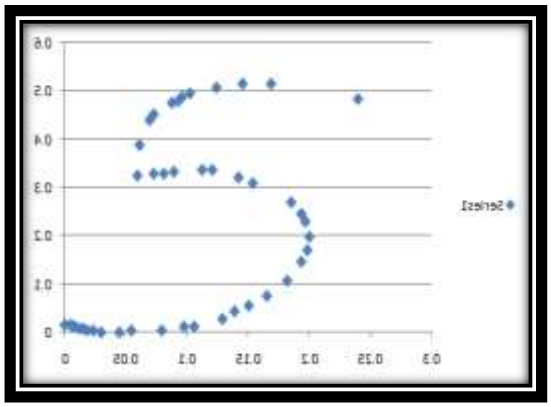
الشكل (7): مخطط مراحل الدراسة

8. اختيار قيمة المعيار لخوارزمية التبسيط

تم اختيار معايير محددة لاستخدامها في عملية تقليص عدد النقاط وإجراء خوارزمية التبسيط على الأرقام المدخلة وهذه المعايير هي: 0.1، 0.01، 0.001. تم اختيار هذه المعايير خصيصا لأنها صغيرة في قيمتها بحيث تتناسب مع حجم الرقم المدخل وفي نفس الوقت تكون متباعدة في مقدار نسبة تقليص النقاط.

أ. **تخزين النقاط الجديدة للأعداد:**

يتم تخزين النقاط الجديدة المكونة للرقم (أي بعد عملية تقليص النقاط) في ملف نصي يكون امتداد هذا الملف نفس امتداد الملف المخزن فيه الكود البرمجي وفي نفس المسار. يتم تخزين هذه النقاط على هيئة (Y،X) والشكل التالي يبين عملية التخزين. ويتم عرض النقاط المكونة للرقم في برنامج Microsoft Excel لكي يسهل معرفة عدد نقاط كل رقم ومن ثم فرزها. وتمت عملية فرز وغربلة الأرقام من خلال عرض شكل الرقم كما هو موضح بالشكل (8).



الشكل(8): عملية فرز الأرقام

ب. **المقارنة بين المعايير:**

تمت المقارنة بين المعايير من خلال:

- الشكل العام للرقم:

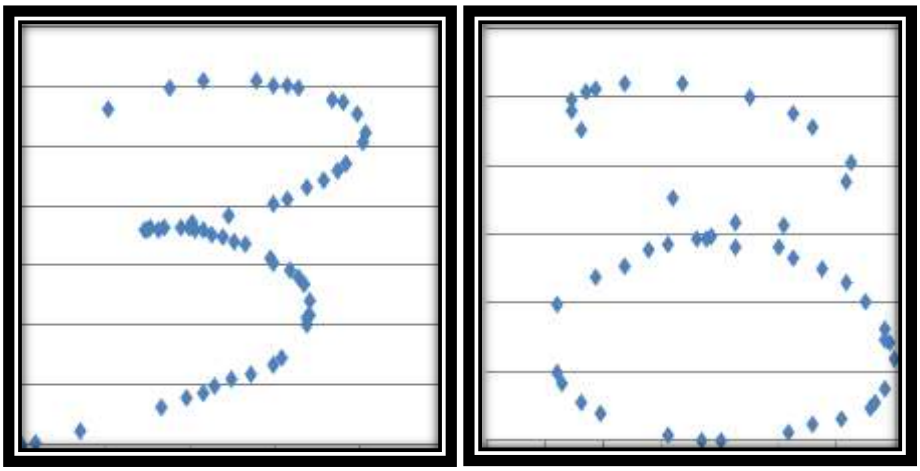
تمت المقارنة بين المعايير المستخدمة في هذه الدراسة من خلال عرض شكل الأرقام في برنامج Microsoft Excel وتم العرض عن طريق تظليل محوري (X،Y) ومن ثم الضغط على قائمة إدراج ومن ثم اختيار الخيار Scatter (مبعثر) وبعدها يتم تحديد الهيئة المراد عرض الرقم بها ومن ثم يظهر مباشرة الشكل النهائي للرقم بعد عملية التقليص وهذه الخطوات تتم على جميع الأرقام.

- تقليص النقاط:

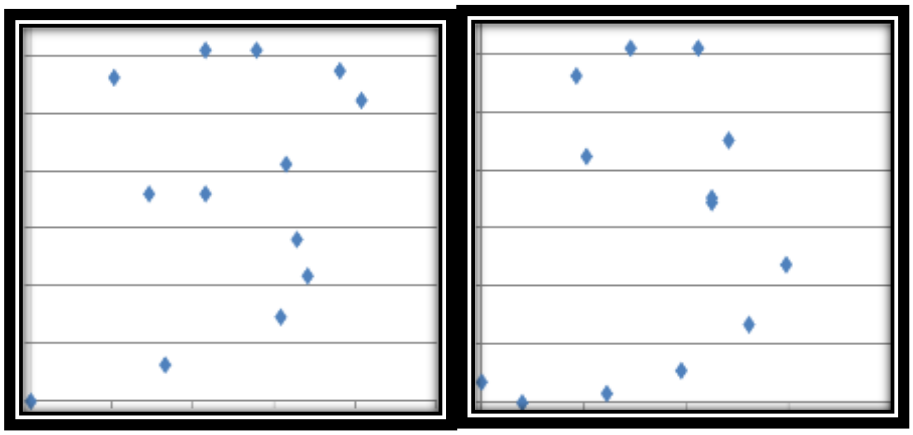
تم تحديد الفروقات في نسبة التقليص بين المعايير من خلال برامج Microsoft Excel وذلك عن طريق فتح المجلد الخاص بكل كاتب وتحديد الملف الخاص بكل رقم والتي من خلالها يتم معرفة عدد النقاط المكونة للرقم المكتوب ويتم تحديد نسبة تقليص النقاط قبل وبعد التقليص.

9. النتائج

تم في هذه الدراسة تجربة ثلاثة معايير مختلفة لمقدار التفاوت في خوارزمية Douglas- Peucker لتقليص وتبسيط عدد النقاط في المنحنيات. تمت مقارنة النتائج المتحصل عليها بطريقتين، وهما نسبة التقليص الحادث من تطبيق الخوارزمية على الأرقام المدخلة. الأشكال (9) و (10) و (11) توضح أشكال الأرقام 8 و 3 بعد إجراء وتطبيق الخوارزمية عليها للمعايير 0.001 و 0.01 و 0.1 على الترتيب.



الشكل (9): شكل الأرقام باستخدام المعيار الأول (0.001)



الشكل (10): شكل الأرقام باستخدام المعيار الثاني (0.01)

الشكل (11): شكل الأرقام باستخدام المعيار الثالث (0.1)

والطريقة الثانية هي مقدار معرفة الشكل الجديد للرقم بعد عملية التقليص باستخدام استبيان معد لهذا الغرض كما هو مبين بالشكل (12).

سيتم استعراض ملخص النتائج المتحصل عليها في كلتا الطريقتين كما يلي:

نسب التقليص للأرقام:

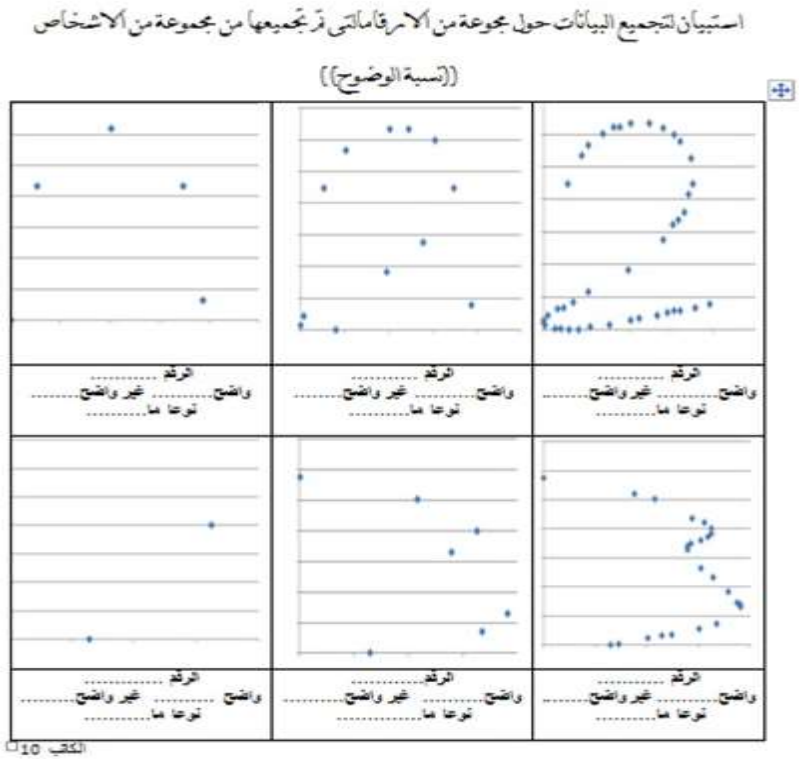
تم تلخيص نسبة التقليص للمعايير الثلاثة لجميع الأرقام المدخلة من 0-9 في الجدول (1) حيث يبين أن المعيار 0.001 قلص عدد النقاط للأرقام بنسب تتراوح بين 40% و63% حسب شكل الرقم. أما المعيار 0.01 فنسبة التقليص كانت بين 77% و 92% تقريبا. والمعيار 0.1 كانت نسبة التقليص مرتفعة بين 91% و 94% تقريبا.

جدول (1): متوسط نسب النقاط المقلّصة لجميع الأرقام للمعايير الثلاثة

الرقم	نسبة المعيار0.001 لكل رقم	نسبة المعيار0.01 لكل رقم	نسبة المعيار 0.1 لكل رقم
1	63.05%	92.08%	94.74%
2	44.19%	83.92%	93.87%
3	40.45%	81.45%	94.06%
4	51.39%	85.29%	92.48%
5	44.86%	82.44%	93.37%
6	41.00%	77.49%	92.72%
7	48.35%	85.62%	91.17%
8	40.88%	78.11%	93.89%
9	45.81%	80.22%	94.36%
0	44.12%	77.31%	92.78%

نسبة الوضوح:

في الطريقة الثانية تم عمل استبيان لمعرفة وضوح الأرقام للمستخدمين بشكل مباشر. حيث تم توزيع الاستبيان كما بالشكل (12) على عدد من الأشخاص لاستبيان مدى وضوح الأرقام بعد عملية التقليص لكل معيار وتم تجميع 48 عينة من هذا الاستبيان المباشر. الجدول (2) يوضح نسبة وضوح العينات التي تم استخدامها في الاستبيان.



الشكل (12): نماذج من الاستبيان الذي تم استخدامه في هذه الدراسة

جدول (2): تحديد نسبة وضوح العينات

المعيار	عدد العينات	نسبة الوضوح					
		واضح	النسبة	نوعا ما	النسبة	غير واضح	النسبة
المعيار الأول	48	48	100%	0	0%	0	0%
المعيار الثاني	48	21	44%	7	15%	20	42%
المعيار الثالث	48	7	15%	0	0%	39	81%

من خلال الجدول (2) من المُلاحظ أن المعيار الأول 0.001 حقق نسبة 100% من درجة الوضوح من خلال التعرف على جميع الأرقام التي تم عرضها في الاستبيان. أما المعيار الثاني 0.01 فكانت نسبة الوضوح 44% و 15% فقط للمعيار الثالث0.1.

10. الاستنتاجات

1- اتضح من النتائج المتحصل عليها من خلال الاستبيان أن المعيار الأول 0.001 هو أفضل معيار بالنسبة للشكل العام وأن المعيار 0.1 هو أفضل معيار بالنسبة للنظام.

2- تم أيضا استنتاج أن العلاقة بين معيار التقليص والنص المُدخل عكسية كلما كان معيار التقليص أكبر كان عدد النقاط المكونة للنص أقل وبالعكس.

3- يمكن استنتاج أنه كلما قل عدد النقاط المكونة للنص كانت عملية المعالجة أسهل وأسرع بالنسبة للنظام وكلما زاد عدد النقاط كان الشكل العام أوضح بالنسبة للمستخدم.

11. المراجع

[1]

Mustafa Ali Abuzaraida, Akram M. Zeki and Ahmed M. Zeki, "The Importance of Preprocessing Phase in Online Arabic Character Recognition System," *International Journal on Islamic Applications in Computer Science And Technology- IJASAT,* vol. 1, pp. 14-23, 2013.

[2]

Mustafa Ali Abuzaraida, Akram M. Zeki and Ahmed M. Zeki, "Problems of Writing on Digital Surfaces in Online Handwriting Recognition Systems," In Proceeding of the 5th International Conference onInformation and Communication Technology for the Muslim World (ICT4M), Rabat, Morocco, 2013, pp. 1-5.

[3]

Mustafa Ali Abuzaraida, Akram M Zeki and Ahmed M Zeki, "Recognition Techniques for Online Arabic Handwriting Recognition Systems," In Proceeding of the International Conference on Advanced Computer Science Applications and Technologies (ACSAT2012), Kuala Lumpur, Malaysia, 2012.

[4]

Mustafa Ali Abuzaraida, Akram M Zeki and Ahmed M Zeki, "Segmentation Techniques for Online Arabic Handwriting Recognition: A survey," In Proceeding of the International Conference on Information and Communication Technology for the Muslim World (ICT4M), Jakarta, Indonesia., 2010, pp. D37-D40.

[5]

Mustafa Ali Abuzaraida, Akram M Zeki and Ahmed M Zeki, "Feature Extraction Techniques of Online Handwriting Arabic Text Recognition," In Proceeding of the 5th International Conference on Information and Communication Technology for the Muslim World (ICT4M), Rabat, Morocco, 2013, pp. 1-7.

[6]

Mustafa Ali Abuzaraida, Akram M Zeki and Ahmed M Zeki, "Online Recognition System for Handwritten Arabic Mathematical Symbols," In Proceeding of the Second International Conference on Advanced Computer Science Applications and Technologies (ACSAT2013), Kuching, Malaysia, 2013.

[7]

Mustafa Ali Abuzaraida, Akram M Zeki and Ahmed M Zeki, "Online Recognition System for Handwritten Arabic Digits," In Proceeding of the The 7th International Conference on Information Technology, Amman, Jordan, 2015, pp. 45-49.

[8]

Mustafa Ali Abuzaraida, Akram M Zeki and Ahmed M Zeki, "Online Database of Quranic Handwritten Words," Journal of Theoretical & Applied Information Technology, vol. 62, 2014.

[9]

Mustafa Ali Abuzaraida, Akram M Zeki, Ahmed M Zeki and Nor Farahidah Za'bah, "Online Recognition System for Handwritten Arabic Chemical Symbols," In Proceeding of the Computer and Communication Engineering (ICCCE), 2014 International Conference on, 2014, pp. 138-141.

[10]

Douglas David and Peucker Thomas, "Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature," *Cartographica: The International Journal for Geographic Information and Geovisualization,* vol. 10, pp. 112-122, 1973.

الخلاصة باللغة الانجليزية

Detecting the Suitable Reduction value of Douglas- Peucker Algorithm in Online Handwritten Recognition Systems

Mustafa Ali Abuzaraida
Computer Science Department
Faculty of Information Technology
Misurata Universuty, Libya

Salma Ebrahim alfitouri, Asma Mohammed Al Samo, Aml Sliman Alftisey, Rukaya Ramadan Abu Ejelah
Computer Department
The Higher Institute For Engineering Professions,
Misurata Libya

Abstract. Online handwritten script recognition systems normally contain four main phases namely: Script acquisition, Preprocessing, Feature extraction, and Recognition phases. The goal of preprocessing phase is to reduce or eliminate imperfections which might cause in the acquisition phase. Preprocessing phase also aims to minimize handwriting variations irrelevant for pattern classification which may exist in the script acquisition. The preprocessing phase has a great influence on subsequent processing, and a real impact on the recognition rate. This phase could include a number of steps like resizing, centering, simplifying, and smoothing the text. In this paper, an experimental study is presented on using a simplification technique. This technique is used in the preprocessing phase in online handwritten text recognition system called Douglas-Peucker algorithm. The proposed system is designed to deal with acquired digits form 0 to 9 which will give a deep understanding of using preprocessing steps in this field.

Keywords: Character Recognition, Digits Recognition, Text Simplification, Douglas-Peucker Algorithm.