



## Use The Positive Selection Algorithm and The Negative Selection Algorithm in The English Character Recognition

**Maha A.M. Hassan**

**Noura Abdel Fattah**

**Maha H. Najam**

**Nour S. Shiban**

College of Computer  
Science and  
Mathematics/  
Department of Software  
Engineering

College of Computer  
Science and  
Mathematics/  
Department of Software  
Engineering

College of Computer  
Science and  
Mathematics/  
Department of Software  
Engineering

College of Computer  
Science and  
Mathematics/  
Department of Software  
Engineering

[mahahassan@gmail.com](mailto:mahahassan@gmail.com)

[nourafattah@gmail.com](mailto:nourafattah@gmail.com)

[mahanajam@gmail.com](mailto:mahanajam@gmail.com)

[nourshiban@gmail.com](mailto:nourshiban@gmail.com)

DOI: [10.33899/edusj.1970.163331](https://doi.org/10.33899/edusj.1970.163331)

**Received**  
**23 / 10 / 2013**

**Accepted**  
**05 / 03 / 2014**

### Abstract

In this paper, I use some artificial intelligence techniques algorithms, which are the algorithms of the artificial immune system (positive selection algorithm and negative selection algorithm). Through these two algorithms, it is distinguished between the original and distorted English letters.

Through the results of the patterns, which were obtained after the implementation of these two algorithms, I observed that the percentage of discrimination was 85% for the positive selection algorithm, and 80% for the negative selection algorithm. This discrimination was applied after the images had been divided into blocks, and then the properties of the images had been extracted by adopting the method of seven undependable momentums in order to gain patterns which are close to the original pattern of the introduced English letter.

**Keywords:** Artificial Immune System (AIS) , Feature Extraction, positive selection algorithm and negative selection algorithm.

## استخدام خوارزمية الانتقاء الايجابي وخوارزمية الانتقاء السلبي في تمييز الحروف الانكليزية

نور سالم شيبان	مها حمادي نجم	نورة عبدالفتاح	مها عبدالاله محمد حسن
كلية علوم الحاسوب والرياضيات/ قسم هندسة البرمجيات	كلية علوم الحاسوب والرياضيات/ قسم هندسة البرمجيات	كلية علوم الحاسوب والرياضيات/ قسم هندسة البرمجيات	كلية علوم الحاسوب والرياضيات/ قسم هندسة البرمجيات
nourshiban@gmail.com	mahanajam@gmail.com	nourafattah@gmail.com	Mahahassan@gmail.com
DOI: <a href="https://doi.org/10.33899/edusj.1970.163331">10.33899/edusj.1970.163331</a>			

القبول

2014 / 03 / 05

الاستلام

2013 / 10 / 23

### الملخص

تم في هذا البحث استخدام بعض من خوارزميات التقنيات الذكائية الاصطناعية وهي خوارزميات النظام المناعي الاصطناعي (خوارزمية الانتقاء الايجابي وخوارزمية الانتقاء السلبي) حيث تم اجراء عملية تمييز بين الحروف الانكليزية الاصلية وبين الحروف الانكليزية المشوهة باستخدام الخوارزميتين المذكورتين .

من خلال نتائج الأنماط التي تم الحصول عليها بعد تنفيذ الخوارزميتين نلاحظ ان نسبة التمييز كانت 85% لخوارزمية الانتقاء الايجابي و80% لخوارزمية الانتقاء السلبي بعد ان تمت عملية تقسيم الصور الى blocks ثم استخلاص خواص الصور عن طريق اعتماد طريقة العزوم السبعة للاعتمادية في الحصول على أنماط مقارنة للنمط الاصيلي للحرف الانكليزي المدخل .

**الكلمات المفتاحية:** نظام المناعة الاصطناعي , استخلاص الخواص , خوارزمية الانتقاء الايجابي وخوارزمية الانتقاء السلبي.

### 1. المقدمة

إن أتمتة الخط كانت وما زالت مجالاً مهماً منذ بدايات الحاسبات. ففي العام 1972 أظهر مسحاً اجري ما يقارب 130 عملاً في هذا المجال [1] , وبالرغم من قدم الموضوع إلا أن تمييز الحروف ما زال واحداً من أكثر المجالات تحدياً وإثارة من مجالات البحث في علوم الحاسبات, إذ تطور هذا المجال في السنوات الأخيرة إلى هيكل ناضج ومنتج لأكثر الأعمال وأفضلها. فعلى الرغم من التنبؤات الطويلة المدى بأن الكتابة اليدوية حتى الورقة نفسها ستلغى بوجود الحاسبات الرقمية إلا أن كليهما ما زال قائمين لحد الآن. في حين سهّل الحاسوب عملية إنتاج الوثائق المطبوعة بشكل كبير, إلا أن الراحة والسهولة التي تمنحهما الورقة والقلم تجعلهما دائماً الوسيلة الطبيعية لإنجاز الكثير من المهام [2] .

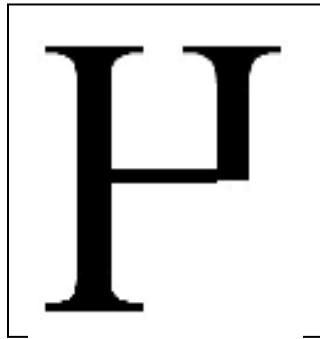
بسبب تطبيقات النظام المناعي AIS الواسعة والفعالة وخصوصا في مجال تمييز الأنماط pattern recognition, فقد تم تنفيذ مرحلة التمييز recognition لنظامنا الحالي باستخدام الخوارزميات المناعية وهي خوارزمية الانتقاء الايجابي وخوارزمية الانتقاء السلبي وذلك لكونها أثبتت فعاليتها في مجال تمييز الأنماط patterns recognition في العديد من الأنظمة السابقة مثل نظام تمييز الأحرف الهندية [3].

## 2. الاعمال السابقة

تكون معظم التقنيات المستخدمة في أنظمة تمييز الخط عبارة عن أنظمة شبكات عصبية [4], إلا أن نظامنا الحالي يختلف عن هذه الأنظمة, وذلك لكون الأنظمة المعتمدة على الشبكات العصبية في التمييز عادة ما تحتاج إلى عدد كبير من النماذج التي تحتاج إليها للتدريب training فضلاً عن الوقت الكبير الذي تحتاجه هذه الأنظمة في التدريب من أجل تحقيق نسبة تمييز جيدة, أما بالنسبة لبحثنا الحالي فقد تمكنا من تحقيق نسبة تمييز عالية مقارنة مع غيره من أنظمة تمييز الخطوط, وفيما يأتي سنعرض عدداً من الأنظمة السابقة المشابهة للعمل المقدم ونقوم بمقارنة نتائجها مع النتيجة التي استطعنا تحقيقها في نظامنا الحالي, [5] الذي اعتمد على استخدام خوارزمية الاختيار النسيلي clonal selection Algorithm المناعية في تمييز الأرقام الهندية فقد وصلت نسبة التمييز فيه إلى 76% وذلك بعد تدريب النظام باستخدام قاعدتين بيانيتين, أما بالنسبة لـ [6] لسنة 2006, الذي يقوم بتمييز الحروف الانكليزية بالاعتماد على أسلوب يجمع ما بين نمطين مختلفين لتطبيع normalization صورة الحرف وطريقة استخلاص خصائص features مهجنة وقوية, ومن ثم يتم التصنيف باستخدام Support Vector Machine, إذ وصلت نسبة التمييز إلى 72,68% , أما في [7] لسنة 2009, تم تنفيذ نظام لتمييز الحروف الانكليزية باستخدام الشبكات العصبية وقد تم تحقيق معدل نسبة تمييز وصل إلى 79,6% باستخدام نماذج تدريب وصلت إلى 592 ونماذج اختبار وصلت إلى 825 نموذجاً.

## 3. هدف البحث

يهدف البحث الى اجراء عملية تمييز صورة الحرف الانكليزي المشوه حاسوبياً (فقدان جزء من الحرف) على سبيل المثال الحرف الانكليزي المشوه (H) الجزء السفلي فيه أي الجزء المفقود يعوض بوحدات فيصبح مشابه لـ Background لصورة الحرف كما في الشكل (1) وتم التمييز باستخدام خوارزميات النظام المناعي الاصطناعي وهي خوارزمية الانتقاء الايجابي وخوارزمية الانتقاء السلبي وقد ساعد في تحسين نسبة التمييز وزيادة سرعة عملية التمييز من خلال تقطيع صورة الحرف الى Blocks واستخلاص خواص الصورة عن طريق حساب العزوم السبعة لقطع الحرف الانكليزي الاصلي والمشوه.

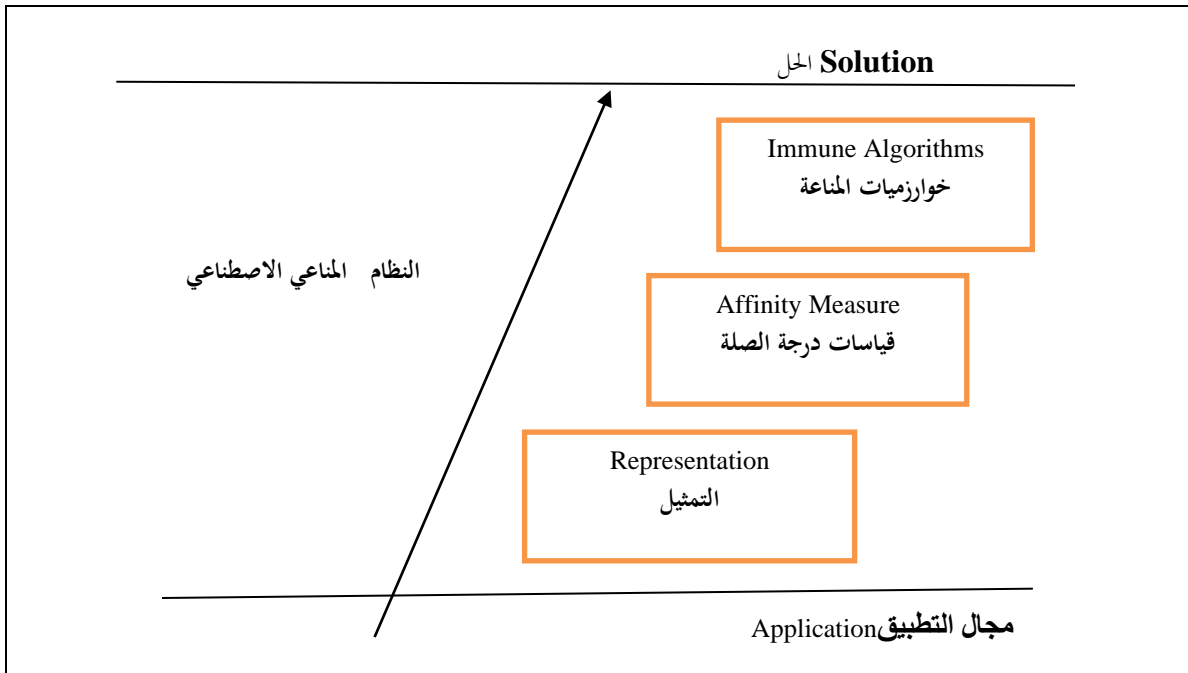


الشكل (1) يمثل صورة الحرف الانكليزي (H) المشوه حاسوبياً

#### 4 . نظام المناعة الاصطناعي (AIS) Artificial Immune System

هو منهج متطور حديث مستوحى من جهاز المناعة الطبيعي في الإنسان, ظهر في التسعينات على أنه فرع جديد من الحساب التطوري Evolutionary computation [8] ان جهاز المناعة الطبيعي معقد جداً لئتم محاكاته صناعياً, ولكن نجح A.B.Watkins بمحاكاة أهم وظائف جهاز المناعة الطبيعي فيما يخص تمييز الأنماط, إن نظام المناعة الاصطناعي هو مجموعة من خوارزميات التصنيف الذكية التي تستخدم آلية دفاع المناعة الطبيعية لأغراض تقنية قادرة على التكيف والتعلم , لذلك انتشر هذا المفهوم في تطبيقات تقنية عديدة خلال العقود الأخيرة [9] , إن الحوسبة المستخدمة أحيائياً (Biologically Inspired Computing) وخصوصاً أنظمة المناعة الاصطناعية (AIS) هي حل واعد لتطوير أدوات دفاعية مكيفة ومؤتمتة للتهديدات الحالية والمستقبلية في اكبر أنظمة (تقنية المعلومات) [1][10].

قدم Timmis و De.Castro إطار عمل مستخدم بصورة أكثر شيوعاً في هندسة AIS يتضمن إطار العمل ثلاث خطوات مستقلة بشكل نسبي في بناء (AIS) أن أساس كل نظام هو مجال التطبيق [11], ويعمل تمثيل البيانات على التعريف بكيفية تحويل بيانات التطبيق الملاحظة إلى بيانات متلائمة النسق وذلك ليتم معالجتها بواسطة خوارزميات (AIS) تحتاج البيانات المستخدمة أن تكون منسقة لتكون قادرة على قياس درجة الصلة في مجال تمثيل البيانات [2] , إن قياس درجة الصلة هو في صلب التصنيف والتمييز بواسطة خوارزميات (AIS) (على الأقل هذا صحيح بالنسبة لخوارزميات (AIS) التكيفية) [9] والشكل (2) يوضح طبقات إطار عمل (AIS).



الشكل (2) يوضح طبقات إطار عمل (AIS) [8]

يملك جهاز المناعة الاصطناعي ( AIS ) عدة صفات حسابية مميزة، وفيما يلي بعض منها [12] [13]:

1. تمييز الأنماط ( Pattern Recognition ): يستطيع جهاز المناعة أن يميز مستضدات معينة ويولد استجابات مناسبة حيث يتم انجاز ذلك بواسطة آلية التمييز المعتمدة على الارتباط للمستقبلات و المستضدات. يعتمد هذا الارتباط على أشكالها الجزيئية و الإشارات [13].
2. استخلاص الخواص ( Feature Extraction ) : بصورة عامة لا ترتبط مستقبلات المناعة بمستضد كامل بل أجزاء منه ووفقا إلى ذلك يمكن لجهاز المناعة أن يميز مستضداً فقط عن طريق التطابق مع أجزاء منه [7].

## 5. استخلاص الخواص Feature Extraction

إنّ عملية استخلاص الخواص واحدة من أهم عناصر نظام تمييز الانماط ، إذ يعتمد أداء نظام التصنيف بشكل مباشر على عملية اختيار الخواص و استخلاصها [6]، الفائدة الرئيسية من استخلاص الخواص هي تقليص حجم البيانات المعروضة بالصورة وتمثل الصورة بمجموعة من الخواص بصيغة رقمية وهذه الخواص تستخدم في نظام التمييز لتصنيف البيانات وتجعل خوارزمية استخلاص الخواص في عملية التمييز أكثر فاعلية و أعلى كفاءة ، إذ أنّ نظام التمييز يتعامل مع مفاهيم رياضية و تقنية من أجل تصنيف مختلف الأجسام (object) من خلال ملاحظة معلومات الصورة مثل المستويات الرمادية للنقط الصورية (pixel) في الصور [4] ومن الحسابات الاحصائية المستخدمة في استخلاص الخواص هي العزوم اللاعتمادية (Moment invariants) تحسب هذه العزوم باستعمال مجموعة لا خطية من اللحظات الهندسية (Moment geometric) والتي لها الإمكانية المطلوبة بأنها ثابتة (invariant) بعد ازاحة الصورة او تكبيرها او تصغيرها او تدويرها وهناك مجموعة عزوم مركزية (Central Moments) لصورة ثنائية البعد ممكن إيجادها بخطوات عدة:

الخطوة الأولى : حساب العزوم الثلاثة الأولى من المعادلة الآتية: [11]

$$m_{pq} = \sum \sum x^p y^q f(x, y) \quad \dots\dots\dots(1)$$

اذ ان:

$M_{pq}$  : هي سلسلة من العزوم ثنائية البعد.

$1, 0 : p, q$

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \text{And} \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad \dots\dots\dots(2)$$

إذ إن:

$\bar{x}$ : مركز الصورة عند الإحداثي السيني.

$\bar{y}$ : مركز الصورة عند الإحداثي الصادي.

الخطوة الثانية: تحسب العزوم المركزية كما في المعادلة الآتية:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad \dots\dots\dots(3)$$

إذ إن:

$\mu_{pq}$ : هي سلسلة من العزوم ثنائية البعد.

$p, q : 0, 1, 2, 3, \dots$

$$\mu_{00} = \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^0 f(x, y) = \sum_x \sum_y f(x, y) = m_{00}$$

$$\mu_{10} = \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^0 f(x, y) = m_{10} - \frac{m_{10}}{m_{00}} (m_{00}) = 0$$

$$\mu_{01} = \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^1 f(x, y) = m_{01} - \frac{m_{01}}{m_{00}} (m_{00}) = 0$$

$$\mu_{11} = \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^1 f(x, y) = m_{11} - \frac{m_{10}m_{01}}{m_{00}} = m_{11} - \bar{y}m_{10}$$

$$\mu_{20} = \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^0 f(x, y) = m_{20} - \frac{2m_{10}^2}{m_{00}} + \frac{m_{10}^2}{m_{00}} = m_{20} - \frac{m_{10}^2}{m_{00}} = m_{20} - \bar{x}m_{10}$$

$$\mu_{02} = \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^2 f(x, y) = m_{02} - \frac{m_{01}^2}{m_{00}} = m_{02} - \bar{y}m_{01}$$

$$\mu_{21} = \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^1 f(x, y) = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01}$$

$$\mu_{12} = \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^2 f(x, y) = m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10}$$

$$\mu_{30} = \sum_x \sum_y (x - \bar{x})^3 (y - \bar{y})^0 f(x, y) = m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2m_{10}$$

$$\mu_{03} = \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^3 f(x, y) = m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01}$$

والعزوم المركزية الطبيعية (Normalized Central Moments) تحسب كما في المعادلة الآتية:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^r} \dots \dots \dots (4)$$

إذ إن:

$$r = \frac{p + q}{2} + 1 \dots \dots \dots (5)$$

من المعادلات الآتية تحسب العزوم السبعة بالاعتماد على قيم العزوم المركزية الطبيعية :

$$\begin{aligned}
 \phi_1 &= \eta_{20} + \eta_{02} \\
 \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
 \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
 \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
 \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
 &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
 \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
 \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
 &\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
 \end{aligned}$$

بعد ان يتم تحليل صورة الحرف الانكليزي الاصلي والمشوه من خلال استخلاص خواص صورة الحرف عن طريق حساب العزوم السبعة اللاعتمادية لصورة كل حرف انكليزي حيث ان هذه الطريقة تعد من الطرائق المرغوب فيها وذلك لقدرتها على اخضاع الصورة الى التدوير والازاحة والتكبير ثم بعد ذلك يتم تنفيذ اما خوارزمية الانتقاء الايجابي او خوارزمية الانتقاء السلبي على بيانات صورة الحرف الانكليزي لغرض تمييز هذه الحروف .

## 6. خوارزمية الانتقاء الايجابي وخوارزمية الانتقاء السلبي

### positive selection algorithm and negative selection algorithm

في عام 1992 [14] اقترح نموذج اتمته خلوية لإنجاز محاكاة حاسوبية للجهاز المناعي الاصطناعي سميت بخوارزمية الانتقاء الايجابي حيث تمتلك الخوارزمية خصائص مفيدة من منظور حسابي، الخطوات الاساسية للخوارزمية هي:

المُدخل:  $T_{seen}$  = مجموعة من العناصر الذاتية المعلومة.

المُخرج:  $G$  = مجموعة المستكشفات المتولدة.

أبدأ (begin)

كرر (repeat)

- وُلد مستكشفات بشكل عشوائي وضعها في مجموعة  $P$

- حدد دالة اللياقة لكل عنصر (عضو) في  $P$  مع كل عنصر في المجموعة الذاتية  $T_{seen}$

- اذا استطاع عنصر واحد على الاقل في  $T_{seen}$  التعرف على مستكشف في  $P$  طبقاً لعتبة Threshold، فإن المستكشف يُقبل، والا فانه يضاف الى مجموعة من المستكشفات المتوفرة  $G$  .

( until ) الى ان تتحقق معايير التوقف.

انتهى (end)

تتمثل نقطة بداية الخوارزمية بتكوين مجموعة من الجمل الذاتية والمتمثلة ب T ثم يتم تكوين مجموعة من الكاشفات G والتي تقوم بتمييز متم T فقط ثم يتم تطبيق هذه الكاشفات على بيانات جديدة من اجل تمييزها، في عملنا المقدم فان الرمز T<sub>seen</sub> في الخوارزمية يقابل الرمز C وان الرمز G يقابل الرمز M .

تعد نظرية الانتقاء السلبي فكرة مفيدة جداً لتطوير الخوارزميات التي تراقب الانظمة ضد وقوع حالة الشذوذ او السلوك الغير مألوف [8], ان السمة المفيدة لهذه الخوارزمية هي امكانية استعمالها في اداء المهمات مثل تمييز الانماط وتمييز الصور والامثلية، الخطوات الاساسية للخوارزمية هي:

المُدخل: S<sub>seen</sub> = مجموعة من العناصر الذاتية المعلومة.

المُخرج: D = مجموعة المستكشفات المتولدة.

أبدأ(begin)

كرر(repeat)

- وُلد مستكشفات بشكل عشوائي وضعها في مجموعة P

- حدد دالة اللياقة لكل عنصر (عضو) في P مع كل عنصر في المجموعة الذاتية S<sub>seen</sub>

- اذا استطاع عنصر واحد على الاقل في S<sub>seen</sub> التعرف على مستكشف في P طبقاً لعتبة Threshold، فأن المستكشف يُرفض، والا فانه يضاف الى مجموعة من المستكشفات المتوفرة D .

( until ) الى ان تتحقق معايير التوقف.

انتهى(end)

تتمثل نقطة بداية الخوارزمية بتكوين مجموعة من الجمل الذاتية والمتمثلة ب S ثم يتم تكوين مجموعة من الكاشفات D والتي تقوم بتمييز متم S فقط ثم يتم تطبيق هذه الكاشفات على بيانات جديدة من اجل تمييزها، في عملنا المقدم فان الرمز S<sub>seen</sub> في الخوارزمية يقابل الرمز C وان الرمز D يقابل الرمز M .

لحساب قيمة Affinity التي تمثل درجة الصلة بين صورة الحرف الانكليزي الاصلي والحرف الانكليزي المشوه لكلا الخوارزميتين تحسب بالشكل التالي:

$$\text{Affinity} = \frac{X * 100}{Y} \dots\dots\dots (6)$$

حيث ان:

X : تمثل درجة الصلة بين عنصر في المخزن (P) مع عنصر في المجموعة الذاتية (self set T<sub>seen</sub>) او (self set S<sub>seen</sub>) حسب نوع الخوارزمية المستخدمة.

Y : تمثل عدد صور الحروف الانكليزية المشوهة المدخلة في النظام حيث ان هذه القيمة قابلة للتغيير حسب حاجة المستخدم للنظام .

#### 7. الجانب العملي للبحث

يعتبر النظام المناعي الاصطناعي احدى اهم التقنيات الذكائية الاصطناعية المستخدمة في مجال تمييز الانماط واستخلاص الخواص حيث تم في البحث استخدام خوارزمية الانتقاء السلبي وخوارزمية الانتقاء الايجابي في عملية تمييز انماط الحروف الانكليزية المشوهة علماً ان الحروف الانكليزية الاصلية والمشوهة مكتوبة بخط نوعه (Times New Roman) وحجم الخط (18) حيث تم ادخال الصورة



عن طريق الماسح الضوئي وحفظت الصورة بصيغة (bmp) وبعدها تم تقطيع الصورة وحساب العزوم السبعة ثم اجراء عملية المقارنة عن طريق استخدام خوارزمية الانتقاء السلبي او خوارزمية الانتقاء الايجابي وتم استخدام لغة MATLAB في برمجة النظام، علماً انه لا توجد اية معالجة اولية للأحرف الانكليزية المشوهة المدخلة ، حيث ان الخطوات العملية للبحث تتضمن:

الخطوة الاولى: ادخال مجموعة من الاحرف الانكليزية الصحيحة والمشوهة عن طريق الماسح الضوئي (Scanner).

الخطوة الثانية: حفظ الصور المدخلة بصيغة (bmp) وبحجم (100\*100) وحدة صورية (pixel).

الخطوة الثالثة: يتم تقطيع الصور المدخلة الى عدة (Blocks).

الخطوة الرابعة: يتم حساب العزوم السبعة لكل (Block) عن طريق استخلاص خواص الصور.

الخطوة الخامسة : ثم يتم اختيار العمل باستخدام اما خوارزمية الانتقاء السلبي او خوارزمية الانتقاء الايجابي, فاذا تم اختيار خوارزمية الانتقاء الايجابي فيتم اتباع الخطوات التالية في البرنامج:

الخطوة الاولى: (C) تمثل مجموعة من العناصر مرشحة عشوائياً، (P) تمثل مجموعة الانماط التي توفر الحماية , (M) تمثل مجموعة الكاشفات التي تكون مسؤولة عن تمييز كل عنصر لا يعود للمجموعة (P).

الخطوة الثانية: المقارنة بين العناصر في (C) والعناصر في (P) من ناحية مقدار الصلة فيما بينهم:

• اذا استطاع عنصر من (P) تمييز عنصر من (C) من خلال تنفيذ معادلة الـ Affinity الخاصة بخوارزمية الانتقاء الايجابي وهي كالتالي:

$$\text{Affinity} = \frac{X * 100}{64} \dots\dots\dots (7)$$

حيث ان :

X: تمثل درجة الصلة بين العنصر في المخزن (P) مع عنصر في المجموعة الذاتية (T<sub>seen</sub>).

64: تمثل عدد صور الحروف الانكليزية المشوهة المدخلة في النظام المستخدم .

- اذا كانت قيمة الـ Affinity الناتجة اكبر او مساوٍ لحد العتبة Threshold (قيمة حد العتبة تحدد حسب صورة كل حرف انكليزي اصلي) ، فعندئذٍ يتم تمييز العنصر واختياره بشكل ايجابي ثم خزن العنصر (C) داخل مجموعة الكاشفات (M).
- والا يتم حذف العنصر (C) .

وهكذا تتكرر الخطوات حتى يتم التمييز بين العناصر .

اما اذا تم اختيار خوارزمية الانتقاء السلبي فيتم اتباع الخطوات التالية:

الخطوة الاولى: (C) تمثل مجموعة من العناصر مرشحة عشوائياً، (P) تمثل مجموعة الانماط التي توفر الحماية , (M) تمثل مجموعة الكاشفات التي تكون مسؤولة عن تمييز كل عنصر لا يعود للمجموعة (P).

الخطوة الثانية: المقارنة بين العناصر في (C) والعناصر في (P) من ناحية مقدار الصلة فيما بينهم:

• اذا استطاع عنصر من (P) تمييز عنصر من (C) من خلال تنفيذ معادلة الـ Affinity الخاصة بخوارزمية الانتقاء السلبي وهي كالتالي:

$$\text{Affinity} = \frac{X*100}{64} \dots\dots\dots (8)$$

حيث ان :

X: تمثل درجة الصلة بين العنصر في المخزن (P) مع عنصر في المجموعة الذاتية (S<sub>seen</sub>).

64: تمثل عدد صور الحروف الانكليزية المشوهة المدخلة في النظام المستخدم .

- اذا كانت قيمة ال Affinity الناتجة اكبر من حد العتبة Threshold ( يتم تحديدها حسب صورة كل حرف انكليزي اصلي) ، فعندئذ يتم تمييز العنصر واختياره لعدم تطابق صورة الحرف الانكليزي المشوه مع صورة الحرف الانكليزي الاصلي ثم خزن العنصر (C) داخل مجموعة الكاشفات (M).
  - والا يتم حذف العنصر (C) .
- وهكذا تتكرر الخطوات حتى يتم التمييز بين العناصر .

## 8. مناقشة النتائج

من خلال تطبيق النظام لغرض تمييز الحروف الانكليزية المشوهة وعلى وفق النتائج التي تم الحصول عليها، تم التوصل إلى الاستنتاجات الآتية:

1. بعد دراسة الخوارزميات الموجودة في حقل أنظمة AIS تم اختيار خوارزمية الاختيار الانتقاء الايجابي وخوارزمية الانتقاء السلبي لبناء الحل للمشكلة المدروسة لما فيها من قابليات مناسبة لاداء بعض المهام مثل التعلم باستخدام الحاسبة وتمييز الانماط وتمييز الصور والامتلية ومشاكل كشف تطفل الشبكة والحاسوب .

2. تم تحقيق نسب التمييز الآتية:

نسبة التمييز	نوع الإدخال
بلغت النسبة % 85	في حالة تمييز صور الحروف الإنكليزية المشوهة باستخدام خوارزمية الانتقاء الايجابي
بلغت النسبة % 80	في حالة تمييز صور الحروف الإنكليزية المشوهة باستخدام خوارزمية الانتقاء السلبي

3. تم تكوين قاعدة بيانات تضم صور للحروف الانكليزية المشوهة تحتوي على 64 نموذج بالإضافة الى 26 نموذج من صور للحروف الانكليزية الاصلية ، وحيث ان النماذج المشوهة تم تمييزهم باستخدام خوارزمية الانتقاء الايجابي وخوارزمية الانتقاء السلبي وكانت نسب التمييز حسب الجدول اعلاه، بالإضافة الى ان النظام يمتلك الامكانية في ادخال نماذج اضافية في قاعدة البيانات وحسب حاجة المستخدم للنظام.

4. مقارنة مع اعمال سابقة في مجال التمييز نلاحظ ان [6] الذي اعتمد على استخدام خوارزمية الاختيار النسيلي clonal selection المناعية في تمييز الأرقام الهندية فقد وصلت نسبة التمييز فيه إلى 76% وذلك بعد تدريب النظام باستخدام قاعدتين بيانيتين أما في

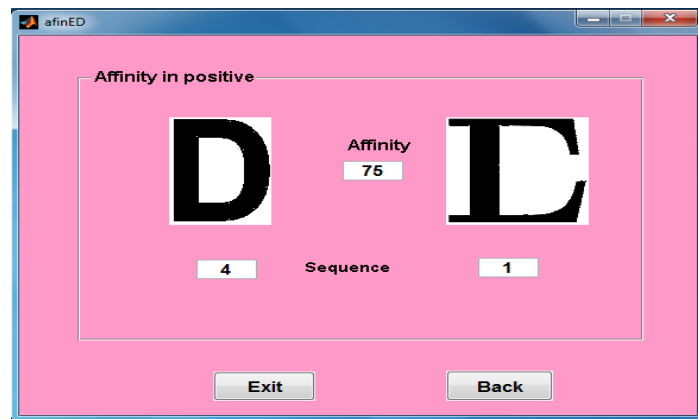
[3] لسنة 2009 تم تنفيذ نظام لتمييز الحروف الانكليزية باستخدام الشبكات العصبية وقد تم تحقيق معدل نسبة تمييز وصلت إلى 79.6% , خلال هذا البحث فقد تم اعتماد حسابات احصائية في استخلاص الخواص لصور الحروف الانكليزية الاصلية والمشوهة من خلال اعتماد طريقة العزوم السبعة للاعتمادية لقدرتها على تعديل الصور من خلال التكبير والازاحة والتدوير, وإن هذه الطريقة مكنة الخوارزميتين من تحسين نسبة التمييز لصور الحروف الانكليزية المشوهة المدخلة للنظام بالاضافة الى مرونة النظام في ادخال عدد صور الحروف الانكليزية المشوهة وحسب حاجة المستخدم للنظام.

### 9. الاستنتاجات

من خلال تطبيق النظام المناعي الاصطناعي لغرض تمييز الحروف الانكليزية وعلى وفق النتائج التي تم الحصول عليها، تم التوصل إلى انه بعد دراسة الخوارزميات الموجودة في أنظمة AIS تم اختيار خوارزمية الانتقاء السلبي negative selection وخوارزمية الانتقاء الايجابي positive selection لما لهما من قابليات عالية في معالجة المعلومات وبنمط عالٍ من التوازي والتوزيع في مجال التصنيف وتمييز الانماط وتمييز الصور.

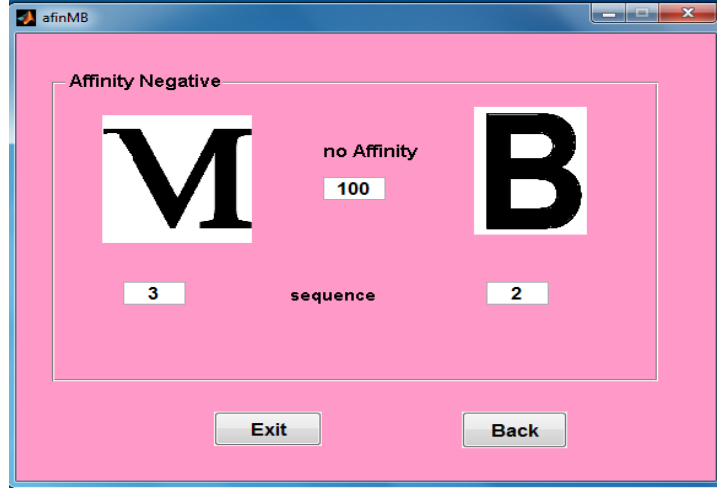
فقد تم في البحث تقطيع كل صورة مدخلة الى (8 blocks) علماً ان حجم (block) الواحد يساوي (144-pixel) وبعد حساب العزوم لكل (blocks) تم الحصول على قيم عزوم كفاءة ومناسبة للتمييز , من خلال النتائج التي تم التوصل اليها نلاحظ انه:

- في خوارزمية الانتقاء الايجابي كلما كانت قيمة الـ Affinity كبيرة لصورة الحرف الانكليزي المشوه مقارنة بقيمة العتبة التابعة لصورة الحرف الانكليزي الاصيلي بمعنى ان صورة هذا الحرف الانكليزي المشوه مقاربة اكثر لصورة الحرف الانكليزي الاصيلي بالنسبة لبقية صور الحروف الانكليزية المشوهة عندئذٍ يتم اختيار هذه الصورة كأفضل صورة ناتجة, كما في الشكل (3).



الشكل (3) يوضح اعلى نسبة تطابق لصورة الحرف الانكليزي المشوه (D) مع صورة الحرف الانكليزي الاصيلي (D) في خوارزمية الانتقاء الايجابي

- في خوارزمية الانتقاء السلبي كلما كانت قيمة ال Affinity كبيرة لصورة الحرف الانكليزي المشوه مقارنة بقيمة العتبة بمعنى ان صورة هذا الحرف الانكليزي المشوه مختلفة (ليست مقاربة) تماماً لصورة الحرف الانكليزي الاصلي بالنسبة لبقية صور الحروف الانكليزية المشوهة عندئذ يتم اختيار هذه الصورة كأفضل صورة ناتجة من الخوارزمية كما في الشكل(4).



الشكل(4) يوضح اعلى نسبة اختلاف(عدم تشابه) لصورة الحرف الانكليزي المشوه (M) مع صورة الحرف الانكليزي الاصلي(B) في خوارزمية الانتقاء السلبي

- إن استخدام خوارزمية الاختيار السلبي والايجابي ساعد كثيرا في تحسين نسبة التمييز حيث ان نسبة التمييز %85 لخوارزمية الانتقاء الايجابي و%80 لخوارزمية الانتقاء السلبي بالاضافة الى زيادة سرعة عملية التمييز عن طريق تقطيع صور الاحرف و استخراج خواص الصورة بحساب العزوم السبعة لقطع الحرف الانكليزي الاصلي والمشوه.

### المصادر

1. Coello, C., Cortez-Rivera, D., Cruz-Cortez, N., (2003), Proc. ICARIS, Springer, pp.1—10.
2. Dasgupta, D., Niño, Luis F., (2009), CRC, ISBN 978-1-4200-6545-9.
3. Goldsby, Richard, A., Kindt, Thomas, J., Osborne, Barbara, A, Kuby, J., (2003), " Immunology ", 5<sup>th</sup> Edition, W. H. Freeman and Company.
4. Andrew Watkins, Xintong BI and Amit Phadke, 2003, Department of Computer Science and Engineering, Mississippi State University, USA, Also: Computing Laboratory, University of Kent, Canterbury, UK.

5. De Castro L.N., Von Zuben, F.J., (2002), IEEE Trans. Evol. Comput., vol. 6, no. 3, pp 239—251.
6. B. Gatos, I. Pratikakis, A.L. Kesidis and S.J. Perantonis, 2006, Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research “Demokritos”, GR-153 10 Agia Paraskevi, Athens, Greece.
7. Divya Sharma, June 2009, Master of Engineering in Software Engineering, Thapar University.
8. Forrest, S., Perelson, A., Allen, L., Cherukuri, R., (1994), Proc. of the IEEE Symposium.
9. De Castro, L. N., Timmis, J., (2002), ISBN 1-85233-594-7, Springer, England.
10. Foster, R. , (2010), PhD thesis, University of Nottingham, England.
11. Haidar Almohri, John S.Gray and Hisham Alnajjar, PhD, 2008.
12. Iqbal, A., (2006), PhD theses, Universiti Teknologi Malaysia.
13. Greensmith, J., Aickelin, U., Cayzer, S., 2005, In ICARIS-05, LNCS 3627, pp. 153–167.
14. Seiden, P.E. and Celada , F.,(1992), J. theor. Bio., 158,pp.329-357.