

تمييز حروف اللغة العربية المكتوبة آلياً باستخدام الشبكات العصبية ذات الانتشار الرجوعي

عاصم عبدالفتاح نبوي، صبري عبدالله محمود *

كلية علوم الحاسب والمعلومات، جامعة الملك سعود، الرياض، المملكة العربية السعودية

*مؤسسة المنارين للتطبيقات الفنية، ص.ب ٥٣٥٣١، الرياض ١١٥٩٣

(قدم للنشر في ١٢/٤/١٩٩٣م، وقبل للنشر في ١٨/٦/١٩٩٥م)

ملخص البحث . تعددت في الآونة الأخيرة البحوث الجارية في حقل التعرف الآلي بطريقة ضوئية على حروف اللغة العربية . وقد كان الباعث على التزايد الملحوظ في وتيرة هذه البحوث عوامل عديدة منها الاهتمام المتزايد بتطبيقات الحاسوب في مجالات جديدة مثل تسيير المكاتب آلياً، والصيرفة، ومكاتب الخدمات البريدية، وكذا التقدم التقني الذي ينزل بكثير من معدات الحاسوب المادية إلى متناول قطاعات أعرض من البشر، بالإضافة إلى التطبيقات الجارية للتعرف على حروف اللغات الأخرى غير العربية .

ويخاطب هذا البحث استخدام الشبكات العصبية ذات الانتشار الرجوعي في التمييز الآلي لحروف اللغة العربية المطبوعة آلياً . وقد تم في هذا البحث استخلاص ملامح الحروف المطلوبة للقيام بالتعرف، مرة باستخدام تحويل " والش " وأخرى باستخدام طريقة العزوم . واستخدمت الشبكات العصبية والمصنفات الإحصائية التقليدية للتعرف على الحروف باستخدام مجموعات مختلفة في مرحلتي التدريب والتعرف، كما تمت المقارنة بينهما . وأجريت مقارنة بين استخدام تحويل " والش " لاستخلاص ملامح الحروف واستخدام طريقة العزوم لنفس الغرض . وقد بينت نتائج البحث أن استخدام تحويل " والش " يؤدي إلى تعرف أفضل وأسرع من استخدام طريقة العزوم، إلا أن الأخيرة تتميز بحاجتها إلى مساحة تخزينية أقل من الأولى . وبينما يتم استخدام المصنفات التقليدية بطريقة مباشرة فإن الشبكة العصبية ذات الانتشار الرجوعي قوية الاحتمال للتشويش وذات

إمكانات واسعة، غير أنها تحتاج إلى بحث مستفيض للوصول إلى التركيب الأمثل المناسب للاستخدام المعين. وقد أثبت البحث أن زيادة مجموعات التدريب أدت إلى تحسين معدلات التعرف في كل الأحوال.

١. مقدمة

يمكن القول بأن الأصول المبكرة لطرائق التعرف على الحروف تعود إلى عام ١٨٧٠م عندما اخترع كاري Carey آلة لمسح الصور ذات شبكية تعتمد على مصفوفة من الخلايا الكهربائية الضوئية لتحويل عناصر الصورة إلى إشارات كهربائية. وقد قام بعد ذلك نيكوف Nipkow عام ١٨٩٠م بابتكار آلة المسح التسلسلية والتي كانت مقدمة لظهور أجهزة بث الصور (التلفزيون) وتجهيزات القراءة الآلية [١].

ولقد ظهرت الحاجة إلى التعرف على الحروف آلياً في البداية كوسيلة لمساعدة المعاقين بصرياً. وتعزى أول محاولة ناجحة في هذا الصدد إلى العالم الروسي تيورين Tyurin في العام ١٩٠٠م. وقد تلت ذلك محاولات فورييه دالب Fourier d'Albe لصناعة الآلة القارئة للحروف الناطقة عام ١٩١٢م، والجهاز التعويضي اللمسي الذي بناه توماس Thomas عام ١٩٢٦م.

ولكن الحقيقة الثابتة أن أسس الطرائق الحديثة لنظم التعرف على الحروف ضوئياً قد تبلورت في منتصف الأربعينيات من القرن الحالي، مع ظهور الحواسيب الرقمية، حيث تبين للباحثين في هذا المجال لأول مرة الإمكانيات الهائلة الكاملة في هذا التطبيق كوسيلة فعالة لمعالجة البيانات في عالم الأعمال التجارية. ويعتبر دافيد شبرد David Shepard - مؤسس شركة أبحاث الآلات الذكية - رائداً في بناء وتطوير معدات التعرف على الحروف ضوئياً بطريقة تجارية.

وهناك ثلاثة تطبيقات رئيسة لنظم التعرف على الحروف ضوئياً تميزت خلال السنوات الأخيرة ويندرج تحت كل منها فروع أخرى. وهذه التطبيقات الثلاثة هي إدخال البيانات المكتوبة ويختص بالمصارف المالية، إدخال النصوص المكتوبة ويختص بتسيير المكاتب آلياً، والتشغيل الآلي للعمليات ويختص بمكاتب الخدمات البريدية. وهذه التطبيقات آنفة الذكر، يضمها جميعاً عامل مشترك هو أنها موجهة لأولئك الذين يمكنهم دفع تكلفتها، مثل المؤسسات العامة والخاصة، والمصارف، والوزارات، وهيئات البريد

وكذلك المؤسسات العسكرية . وقد أدت التقانات الحديثة في عالم الإلكترونيات الدقيقة إلى انخفاض تكلفة إنتاج مثل هذه الآلات إلا أنها تظل في غير متناول يد الأفراد الذين يحتاجون إليها . ويحتوي المرجع [٢] على قائمة بالتطبيقات الممكنة لهذا الموضوع . يشتهر التعرف على الحروف باسم التعرف على الحروف ضوئياً وذلك لأن التعرف المقصود يقتصر على الحروف التي تتم معالجتها ضوئياً ولا علاقة له بالحروف المكتوبة بالحبر المغنطيسي والتي يمكن التعرف عليها دون كبير مشقة . إن أهم طرائق التعرف على الحروف ضوئياً هي التعرف على الحروف ذات الأشكال الثابتة وهي الحروف الطباعية غير اليدوية ، والتعرف على الحروف بطريقة الاتصال المباشر بالحاسوب On-line في أثناء الكتابة : وهو خاص بتلك الحروف المفردة ، غير الموصولة ، التي تكتب باليد بقلم خاص موصول بالحاسوب وتتم معالجتها في أثناء الكتابة الأمر الذي لا يزود نظام التعرف بصورة الحرف فقط بل بمعلومات عن توقيت كتابة كل جزء Stroke من الحرف أيضاً ، وأخيراً التعرف على الحروف المكتوبة باليد وهو خاص بالحروف التي تكتب مفردة ، غير موصولة ، ولا جارية ولا مكتوبة بيد خطاط ، والتعرف على الكتابة اليدوية .

ويتضح مما سبق ترتيب طرائق التعرف على الحروف من حيث الصعوبة حيث تتدرج من أصعبها وهو التعرف على الكتابة اليدوية بحروف جارية إلى أسهلها وهو التعرف على الحروف ذات الأشكال الثابتة . ومن الجدير بالذكر أن غالبية طرائق التعرف على الحروف التي تم نشر بحوث عنها حتى الآن قد اقتصر على التعامل مع حروف لغة واحدة ، مثل الإنجليزية (اللاتينية) واليابانية والصينية والسيريلية Cyrillic (الروسية) والعربية والهندية واليونانية . ويعتبر مرجع [٣] من الأمثلة على البحوث التي عالجت نصوصاً ثنائية اللغة . ويمكننا تلخيص منهجيات التعرف على الحروف فيما يلي :

١- المقارنة الشاملة نقطة بنقطة : وتقتضي هذه الطريقة مقارنة جميع عناصر الصورة مع عناصر النموذج المخزن سابقاً .

٢- التحويلات الشاملة : وكمثال عليها نذكر كارهونن - لواف Karhunen-Loeve ، فورييه Fourier ، حساب العزوم وأخيراً التدوير طبقاً لمحور القصور الذاتي الرئيس .

٣- استخلاص الخصائص الموضوعية : وهذه الخصائص مثل نقاط نهاية الحرف ، الزوايا الواضحة أو الأركان ، والوصلات المتفرعة (على شكل حرف T) ، ونقاط

تقاطع أجزاء الحرف مع بعضها. ويمكن استخلاص هذه الخصائص بطرائق عديدة مثل استخدام الأقنعة الثابتة Templates أو تلك ذات الأوضاع المتغيرة. وكثيراً ما يكون ترقيق الحرف Character thinning شرطاً سابقاً لتطبيق أي طريقة من طرائق استخلاص الخصائص السابق ذكرها.

- ٤- البحث عن الخصائص لاستخدامها بالنسبة لخطوط معينة: وهذه الخطوط يمكن أن تكون أشكالاً رباعية أو أن تكون ببساطة خطوطاً رأسية أو أفقية.
- ٥- التحليل باستخدام الانحناءات Curvatures: وهذا يتطلب تتبع المنحنيات، واكتشاف الحلقات والتقرعات Concavities كما يتطلب تحليلاً هندسياً.
- ٦- الطرائق البنيوية Structural methods: وهذه تستلزم تحليل الحرف إلى عناصر تكوينية وكذا على وصف تركيب بنيته Topological description واختصاره في النهاية إلى رسم تخطيطي بياني.

وتعد مسألة اختيار ملامح حروف العربية، التي يتم على أساسها تمييز هذه الحروف آلياً، من أهم المسائل في هذا الصدد، وذلك لأن معدل نجاح تمييز هذه الحروف يتوقف على هذا الاختيار بصورة كبيرة. ولكل حرف من حروف اللغة العربية بعض الملامح التي تميزه عن غيره من حروفها. ومن بين أهم الملامح المستخدمة في هذا التمييز الحلقات والفجوات (الثقوب) والخطوط الرأسية وما شابه ذلك. ويخاطب هذا البحث طريقتين من طرائق استخلاص الملامح المميزة للحروف العربية وهما بالتحديد: تحويل والش Walsh transform وطريقة العزوم Moment method.

٢. نظرة عامة على نظم التعرف على حروف العربية

على العكس من النصوص اللاتينية، تكتب النصوص العربية دائماً بحروف موصولة من اليمين إلى اليسار. وتحتوي اللغة العربية على ثمانية وعشرين حرفاً أساسياً. ولكل من هذه الحروف أشكال عديدة (من اثنين إلى أربعة أشكال) كما أنها تختلف في الحجم (العرض والارتفاع) طبقاً لموقع الحرف في الكلمة. وبالإضافة إلى ما سبق فإن ستة عشر حرفاً من هذه الحروف تصحبها نقطة أو نقطتان أو ثلاث نقاط، تكتب مفردة أو في خط متصل متعرج zigzag وتستخدم هذه النقاط في التمييز بين

الحروف ذات الجزء الأساسي المشترك .

وبسبب الصفات السابق ذكرها فإن الطرائق الفنية المستخدمة في التعرف على الحروف في النصوص اللاتينية غير قابلة للتطبيق مباشرة في حالة النصوص العربية . ويمكن الوصول إلى نفس هذه النتيجة إذا استعرضنا البحوث العلمية الجارية في حقل التعرف على حروف العربية .

اقترح الشيخ وجندي [٤] نظاماً للتعرف على النصوص العربية المطبوعة . يتم طبقاً لهذا البحث أولاً تجزئة Segmentation النص المكتوب بحروف موصولة إلى حروف مفردة وبعد ذلك يتم نقل الحروف إلى طور التعرف حيث تستخدم معاملات متسلسلة "فورييه" لحدود الحرف على محوري (س، ص) كواصفات للحدود . وبالإضافة إلى معاملات "فورييه" ، فقد استخدمت بعض الملامح الخاصة بتركيب البنية مثل الارتفاع والعرض وكذا عدد العناصر السوداء في صورة الحرف ، في تصنيف العلامات المختلفة للضبط بالشكل . وقد بين البحث أن معدل التعرف على الحروف كان ٩٠٪ بينما كان معدل الرفض ٥,٠٪ ، ومع ذلك فإن نظام التعرف على الحروف في هذا البحث يعتمد على شكل حروف الكتابة الطباعية بسبب اعتماده على الملامح الخاصة بتركيب بنية الحرف .

أما الخالي وسيد أحمد [٥] فقد قدماً نظاماً للتعرف على الحروف في النصوص العربية المطبوعة آلياً . وقد طوراً في هذا الصدد خوارزمية فصل تستخدم لتجزئة الكلمات الجارية (الموصولة) إلى حروفها المفردة . ويتم بعد ذلك التعرف على الحروف المجزأة باستخدام واصفات العزوم الثابتة ، وقد أفاد الباحثان أن معدل التعرف كان ٩٠٪ . أما الدابي وآخرون [٦] فقد استخدموا أيضاً العزوم الثابتة التراكمية كواصفات لملامح الحروف . ويسبق مرحلة التعرف - في هذا النظام - مرحلة التجزئة التي يتم فيها تحويل النص ذي الكلمات الجارية إلى سلسلة من الحروف المعزولة . وفي هذه الطريقة (والتي تعتمد أيضاً على شكل حروف الطباعة) استخدم الباحثون عرض أصغر حرف في مجموعة الحروف بوصفه عرضاً أدنى لحروف المجموعة . ويتم بعد ذلك حساب العزوم وفحصها بالمقارنة مع عناصر المتجه الذي يحوي ملامح الحروف . وفي حالة فشل النظام في التعرف على الحرف (بسبب العرض الأدنى) ، يتم تعديل العرض بإضافة عمود آخر

وإعادة حساب العزوم ومقارنتها، وتكرر هذه العملية حتى يتم التعرف على الحرف . وقد قرر الباحثون أن نسبة التعرف كانت ٩٤٪ . ومن الواضح أن طريقة العزوم الثابتة التي استخدمت في هذين البحثين تتطلب حسابات آلية مكثفة، كما أن العزوم الثابتة المستخدمة فيها تعاني من حساسية مفرطة لأقل تغير في الشكل المدخل (للتعرف عليه) كأن تتغير ثخانة جسم الحرف مثلاً.

وقد طور المعلم وياماغوتشي [٧] طريقة للتعرف على النصوص العربية الجارية المكتوبة باليد . تجزأ الكلمات أولاً إلى أجزاء (أشواط) يتم بعد ذلك تصنيفها باستخدام خصائصها الهندسية وتركيب بنيتها، مثل طول الشوط ومكانه، ونقطة الاتصال بينه وبين الشوط السابق وزاوية الاتجاه . وفي النهاية يتم في خطوات عديدة دمج هذه الأشواط في سلسلة من الحروف التي تمثل الكلمة التي تم التعرف عليها . ويقرر الباحثان أن معدل التعرف كان ٩١٪ . ولقد عزا الباحثان فشل النظام في التعرف على بعض الحروف إلى أخطاء تقع في عمليتي تجزئة الكلمات وتصنيف الأشواط .

واستخدم الشيخ والطويل [٨-١٠] طريقة فنية مشابهة تم فيها استخلاص ملامح حروف العربية المعزولة (غير الجارية) المكتوبة باليد بطريقة الاتصال المباشر مع نظام التعرف . وقد بين الباحثان أن متوسط معدل التعرف كان ٩٩٪ ، ومرة أخرى يعاني هذا النظام من كونه معتمداً على شكل الحروف ومما يدل على ذلك أن طور التعرف يعتمد على عدة عتبات ذات ضبط سابق Preset thresholds .

أما الجويلي وآخرون [١١] فقد اقترحوا نظاماً للتعرف على نصوص عربية ضوئية متعددة أشكال الحروف . وينقسم النظام المقترح إلى ثلاثة أطوار . فينتج طور التجزئة- من النص الموصول - مجموعة مبدئية من الحروف وذلك طبقاً لمجموعة من القواعد المحددة سابقاً . ويتم إمرار هذه المخرجات إلى طور تصنيف تمهيدي يتم فيه تصنيف الحروف المجهولة إلى واحدة من عشر فئات ممكنة طبقاً لمجموعة من القواعد المعتمدة على قيم متغيرات خاصة بالحروف تم الحصول عليها في أثناء التدريب . وهذه المتغيرات هي، على سبيل المثال، الارتفاع والعرض وموضع الحرف بالنسبة لخط الأساس Baseline وكذا وجود النقاط، ويعتمد التعرف على الحرف على الملامح الهندسية لحروف اللغة العربية . تم اختبار النظام باستخدام أشكال عديدة للحروف وكان معدل التعرف

٩٤٪. ويعتمد هذا النظام أساساً على شكل معين للحروف بسبب اعتماد مرحلة التصنيف على ارتفاع الحروف وعرضها.

أما أمين والقدافي [١٢] فقد اقترحا طريقة للتعرف الآلي على النصوص العربية المكتوبة بأشكال متعددة للحروف. وتعتمد تجزئة الكلمات هنا على إيجاد النقاط في الرسم البياني النسيجي Histogram الرأسي للكلمة والتي يكون عندها مجموع العناصر السوداء أقل من القيمة المتوسطة التي تم حسابها سابقاً لمجموع الأعمدة. ويصير التعرف بعدئذ على الحروف المفردة بناء على تحليل كل حرف إلى واحد أو أكثر من النماذج الفرعية Sub-patterns. ويقوم النظام بعد ذلك ببناء رسم بياني نسيجي لكل نموذج فرعي لكي يستخرج منه البدائيات Primitives (مثل التوجه الأفقي والرأسي، والاتصال بين النماذج الفرعية باستخدام منحنيات مفتوحة أو مغلقة). وقد قرر الباحثان أن معدل التعرف كان ٩٥,٥٪. ويعاني النظام من عوامل تحد من كفايته في مرحلتي التجزئة والتعرف. وبعبارة أدق، فإن مرحلة التجزئة تتطلب معرفة سابقة بعرض حرف العربية المطلوب التعرف عليه، أما مرحلة التعرف فيعتمد نجاحها على عدد من العتبات السابقة الضبط. ولقد استخدم الباحث الأول طريقة التعرف نفسها في بحثين آخرين [١٣ - ١٤]. قام غورين وآخرون [١٥] بإنجاز نظام تعرف على النصوص العربية، يتم فيه تجزئة النص إلى حروف مفردة. وتعتمد خوارزمية التجزئة على تقسيم الكلمة إلى أشواط أساسية وأخرى ثانوية. يتم بعد ذلك تصنيف الحروف طبقاً لموقع الشوط (جزء الحرف) وشكله ووجود الحلقات فيه. ولقد تم اختبار النظام باستخدام نصوص مطبوعة وأخرى مكتوبة باليد، وكانت النتائج كالتالي: معدل تعرف ٩٢٪ على النصوص المطبوعة، معدل تعرف ٩٠٪ على النصوص المكتوبة باليد. ويستلزم هذا النظام ترقيق النص قبل التعرف، مما يؤدي بالتالي إلى استنفاد وقت كبير في معالجة النص بالإضافة إلى انخفاض دقة التعرف بسبب عدم تفرد Non-uniqueness الصور المرققة Thinned للحرف المطلوب التعرف عليه.

٣. خصائص عامة للنصوص العربية

١ - تكتب كلمات اللغة العربية بحروف موصولة (وتشبه في ذلك الكتابة اللاتينية الموصولة)، وتكتب من اليمين إلى الشمال، وعلى وجه العموم فإن الكلمة العربية

تتكون من جزء أو أكثر من الأجزاء المتصلة، يتكون كل منها من حرف أو أكثر، ويصعب فصل الكلمات الموصولة إلى حروف منفصلة. ويحدث انفصال أجزاء الكلمة العربية نتيجة لوجود الحروف غير القابلة للوصل - من الناحية اليسرى - مع الحروف التالية لها.

٢- يمكن لحرف العربية أن يتخذ أشكالاً عديدة اعتماداً على مكانه في الكلمة (بداية -

نهاية - وسط - مفرد)، مما يزيد من صعوبة التعرف على النصوص العربية.

٣- تتغير حروف العربية طبقاً لأحجامها، حتى لو استخدمنا شكلاً واحداً من أشكال الحروف الطباعية في طباعة النص.

٤- كثير من حروف العربية لها نقاط توضع على مسافات مناسبة فوق الحرف أو أسفله. وتكون هذه النقاط مفردة أو ثنائية أو ثلاثية. وهناك حروف عربية مختلفة لها نفس جسم الحرف وتختلف فقط في عدد ومكان وضع النقاط التي تميز بعض هذه الحروف عن بعضها الآخر.

٥- تستخدم اللغة العربية نوعاً آخر من الحروف الخاصة لضبط الحروف بالشكل، وهي المعروفة بحركات الضمة والفتحة والكسرة بالإضافة إلى الشدة. وعلى الرغم من أن ضبط نفس حرف العربية في كلمة ما باستخدام حركة مختلفة يمكن أن يؤدي إلى كلمة جديدة ذات معنى مختلف، إلا أن قارئ العربية يمكنه غالباً أن يعرف معنى الكلمة المطلوبة من السياق دون حاجة إلى هذه الحروف الزائدة، أي حركات الضبط بالشكل. وعندما تستخدم حركات الضبط بالشكل هذه فإنها تظهر فوق أو تحت الحروف ويتم اعتبارها كيانات مستقلة. هذا بالإضافة إلى أن بعض الحروف العربية مثل الهمزة (ء) والمدة (آ) تستخدم لتعديل نطق بعض حروف العربية، وفي مثل هذه الحالات تظهر بعض هذه العلامات فوق أو أسفل الحروف المعينة بمسافة مناسبة.

٦- بعض حروف العربية قد يتراكب مع الحروف المجاورة له. ويمكن لمدى هذا التراكب أن يتغير طبقاً لشكل حروف الطباعة أو نوعها، أو لتصميم الآلة الكاتبة، أو حتى لعادات الكاتب إن كانت الكتابة يدوية.

٤ . استخلاص ملامح حروف العربية باستخدام تحويل "والش"

يستخدم تحويل "والش" في التطبيقات الخاصة بمعالجة الصور [١٦] . والصيغة العامة لهذا التحويل $W(u)$ عند تطبيقه على دالة $f(x)$ طولها L هي

$$W(u) = \sum_{x=0}^{L-1} g(x,u) f(x) \quad (1)$$

حيث $g(x,u)$ هو الجوهر kernel وهو عبارة عن مجموعة من دوال "والش" . وترتب هذه الدوال على شكل أزواج من الدوال التي تتميز بوجود تماثل فردي وزوجي مع وجود تزايد في معدل تغير إشاراتها :
 L هو طول الدالة $f(x)$ ،
 u هو دليل Index دوال "والش" .

ويمكن تمثيل الجوهر المستخدم $g(x,u)$ بالمعادلة

$$g(x,u) = \frac{1}{N} \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)} \quad (2)$$

$$n = \lceil \log_2(L) \rceil \quad \text{حيث}$$

والعلامة $\lceil \dots \rceil$ تعني أصغر عدد صحيح بحيث تكون $N=2^n \geq L$ ، $b_k(z)$ هي الخانة الثنائية رقم k في التمثيل الثنائي لقيمة z . وباستخدام هذا الجوهر يمكن تحويل الدالة $f(x)$ من عدد من العناصر L إلى عدد آخر N وذلك بإضافة أصفار .

ويمكن التعبير عن تحويل "والش" الرقمي $W(u)$ لدالة $f(x)$ ذات عدد N عنصراً كالتالي :

$$W(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)} \quad (3)$$

ومن الممكن أيضاً حساب تحويل "والش" المنفصل باستخدام خوارزمية سريعة

[١٥]. ومن المعروف أن عدد ملامح الصورة اللازم لتحويل بياناتها إلى رموز يعتمد على دقة التطبيق المطلوب. ومن ناحية أخرى، فإن زمن الحساب الآلي المطلوب للتطبيق المعين يتناسب طردياً مع عدد ملامح "والش" المستخدمة. فمثلاً عندما يكون التطبيق المطلوب هو تمييز الأشكال في عمليات معالجة الخرائط والصور الجوية، تلزم المحافظة على جميع التفاصيل، مما نحتاج معه إلى قدر زائد من تحليل الصورة بطريقة تفصيلية، وتكون هناك بالضرورة حاجة إلى المزيد من الدقة التي تؤدي بدورها إلى الحاجة إلى زيادة العدد المطلوب من ملامح "والش" بصورة كبيرة نسبياً. أما عندما يكون التطبيق هو تمييز الحروف، فإن الدقة في مثل هذا التطبيق لا تعتمد على كل تفاصيل الصورة، فيمكننا استخدام عدد أقل من ملامح "والش" لكي نرمز بيانات الحروف. ولقد تمت في هذا البحث تجربة عدد من المجموعات المؤتلفة من هذه الملامح وتم التوصل إلى نتائج ممتازة بشأن كل من هذين المطلبين المتعارضين: زيادة دقة تمييز الحروف، وتقليل زمن الحسابات الآلية اللازمة لهذا التمييز. ولقد وجد - في هذا الصدد - أن عشرة من ملامح "والش" لكل إسقاط - رأسي أو أفقي - للحرف تكفي للتمييز. ولقد تم في هذا البحث إدخال بيانات الرسم البياني النسيجي الخاصة بالصفوف (الأفقية)، ثم مثلتها الخاصة بالأعمدة (الرأسية)، وتم استخلاص الملامح العشرين المطلوبة لكل حرف من حروف العربية على حدة. ومن الجدير بالذكر أن إسقاط الحرف على محور واحد فقط - رأسيًا كان أو أفقيًا - لا يكون فريداً Unique، ولذا لزم الإسقاط على كلا المحورين. وقد استخدم إسقاط الحرف لتقليل حجم البيانات المعالجة بدرجة كبيرة.

٥. استخلاص ملامح حروف العربية بطريقة العزوم

إن العزوم المحددة Discrete من رتبة $(p+q)$ للدالة المتصلة الثنائية الأبعاد $f(x,y)$ ، والتي نرمز لها بالرمز m_{pq} يمكن تحديدها بالعلاقة الآتية [١٦، ١٧]:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \quad (٤)$$

ويمكن التعبير عن العزوم المركزية Central moments بالصورة الآتية:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (5)$$

حيث

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

ويمكن حساب هذه العزوم المركزية باستخدام دالة العزوم، حيث أن هذه العزوم من الرتب الصفرية والأولى والثانية والثالثة تعطى بالمعادلات الآتية:

$$\begin{aligned} \mu_{00} &= m_{00}, \quad \mu_{01} = 0, \quad \mu_{10} = 0 \\ \mu_{11} &= m_{11} - \bar{y}m_{10}, \quad \mu_{02} = m_{02} - \bar{y}m_{01}, \quad \mu_{20} = m_{20} - \bar{x}m_{10} \\ \mu_{03} &= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01}, \quad \mu_{30} = m_{30} - 3\bar{x}m_{20} + 2m_{10}\bar{x}^2 \\ \mu_{12} &= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10}, \quad \mu_{21} = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01}. \end{aligned} \quad (6)$$

وعند حساب هذه العزوم المركزية لمنطقة ما، R، تكون محل مقارنة مباشرة مع ميكانيكا الأجسام [١٧] كما هو موضح في الجدول التالي:

العزم المركزي	التفسير
μ_{20}	المركزية الأفقية
μ_{02}	المركزية الرأسية
μ_{11}	القطرية: وهي معيار يبين الربع بالنسبة للمركز المتوسط الذي يكون فيه للمنطقة R كتلة أكبر
μ_{12}	التباعد الأفقي: وهو معيار يبين الامتداد النسبي للجانب الأيسر من R مقارنةً بالجانب الأيمن
μ_{21}	التباعد الرأسي: وهو معيار يبين الامتداد النسبي للجانب السفلي من R مقارنةً بالجانب العلوي
μ_{30}	اختلال التوازن الأفقي: موقع مركز الثقل بالنسبة للامتداد الأفقي
μ_{03}	اختلال التوازن الرأسي: موقع مركز الثقل بالنسبة للامتداد الرأسي

ويمكن تحديد العزوم المركزية الموحدة Normalized كما يلي:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \dots \dots \dots (V)$$

$$\gamma = \frac{p+q}{2} + 1 \quad \text{حيث}$$

ويمكن اشتقاق مجموعة من سبعة عزوم وذلك من العزوم المركزية الموحدة ذات الرتبين الثانية والثالثة. وتتميز مجموعة العزوم المشتقة هذه بأنها لا تتأثر بتغير الحجم أو النقل (الإزاحة الخطية) أو الدوران. وقد تبين أن هذا العدد من العزوم يكفي للتعرف على الحروف، وهذه العزوم يمكن اشتقاقها كالتالي:

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

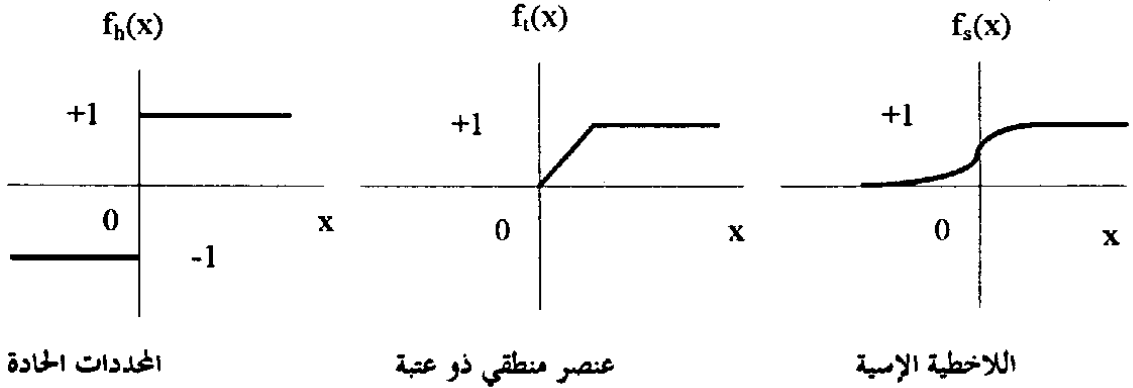
٦ . الشبكات العصبية Neural Networks

إن نماذج الشبكات الكهربائية الاصطناعية الشبيهة بالخلايا العصبية، أو باختصار "الشبكات العصبية"، تعرف بأسماء عديدة مثل النماذج التوصيلية Connectionist models، والنماذج ذات المعالجة الموزعة المتوازية Parallel distributed processing models،

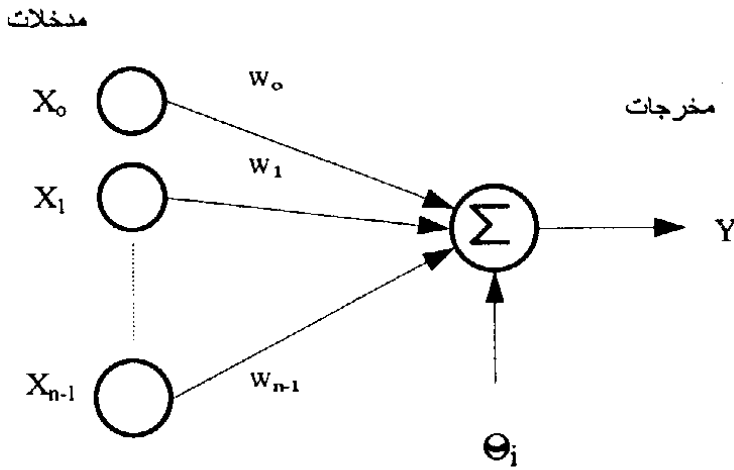
كما تعرف أيضاً باسم النظم ذات الشكل العصبي Neuromorphic systems . ومهما اختلف الاسم الذي نطلقه على النموذج المستخدم، فإن كلاً من هذه النماذج يحاول تحقيق أداء أفضل عن طريق توظيف عناصر بسيطة للحساب الآلي ذات توصيل بيني كثيف . ومن المعلوم أن البنية الاصطناعية للشبكات العصبية مؤسّسة على فهمنا الحالي لكيفية عمل النظم العصبية الحيوية Biological nervous systems . وتجد نماذج الشبكات العصبية أعظم إمكانات استخدامها في حقل التعرف على الصور حيث يلزم اختبار صحة فروض عديدة في آن واحد على التوازي، مما يتطلب قدرات حسابية عالية . وعلى أية حال فإن البون لا يزال شاسعاً بين أداء أفضل نظم الشبكات العصبية المستخدمة حالياً، وبين الأداء البشري المتميز . ويمكن لمن يريد مقدمة عن هذه الشبكات العصبية الرجوع إلى المرجع [١٨] . والعناصر الحسابية - أو العقد Nodes - المستخدمة في نماذج الشبكات العصبية تكون غالباً لاخطية Nonlinear وهي تناظرية Analog بطريقة نموذجية . وأبسط هذه العقد يجمع عدداً من المدخلات ذات أوزان كما هو موضح بالشكل رقم (١) ، ويخضع الناتج لدالة لاخطية كما يلي :

$$y = f \left(\sum_{i=0}^{N-1} W_i x_i + \theta_i \right) \quad (9)$$

وتتميز العقدة بوجود عتبة داخلية (θ) وبنوع الدالة اللاخطية التي تخضع مجموع المدخلات لها قبل إخراج النتيجة . ويبين الشكل رقم (١) ثلاثة أنواع شائعة الاستخدام من هذه الدوال اللاخطية وهي : المحددات الحادة Hard limiters والعناصر المنطقية ذات العتبة Threshold logic elements واللاخطية الإسسية (التي على شكل Sigmoidal) S nonlinearity على الترتيب . ويبين الشكل رقم (٢) عنصراً حسابياً - أو عقدة - يجمع عدد N من المدخلات ثم يخضع الناتج لدالة لاخطية . ويمكن للعقد ذات التركيب الأكثر تعقيداً أن تحتوي على تكامل زمني أو أنواع أخرى من الدوال المعتمدة على الزمن وأن تحتوي أيضاً على عمليات رياضية أكثر تعقيداً من عملية الجمع البسيطة المذكورة آنفاً . ويتم تحديد نموذج الشبكة العصبية باختيار تركيب بنية النموذج وخصائص العقدة وقواعد التدريب والتعليم . وهذه القواعد تحدد مجموعة مبدئية من الأوزان وتبين كيفية تعديل

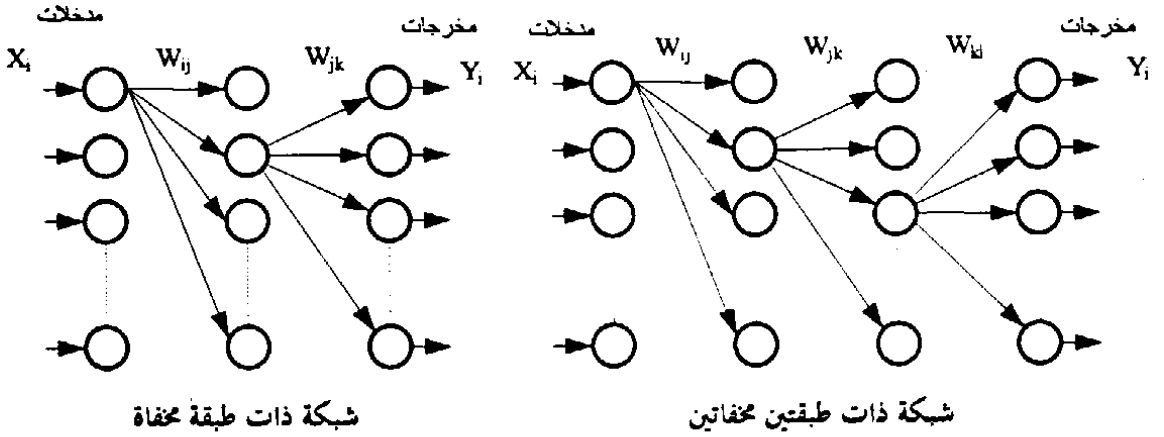


شكل رقم ١ . ثلاثة أنواع من الدوال اللاخطية المستخدمة في الشبكات العصبية .



شكل رقم ٢ . شبكة عصبية بسيطة التركيب .

هذه الأوزان في أثناء استخدام النموذج لتحسين الأداء .
 لقد تم في هذا البحث استخدام شبكة عصبية ذات تركيب بنيوي يسمى بالإدراك المتعدد الطبقات Multi-layer perception . وينتج هذا النوع من التركيب البنيوي شبكة ذات انتشار أمامي تستخدم واحدة أو أكثر من طبقات العقد بين طبقة المدخلات وطبقة المخرجات ، وتسمى هذه الطبقات البينية بالطبقات المخفية Hidden layers . ويبين شكل رقم (٣) نموذجين لشبكتين من الشبكات ذات الإدراك المتعدد الطبقات تستخدم



شكل رقم ٣. شبكتان من ذوات الإدراك المتعدد الطبقات بإحدهما طبقة مخفاة وبالأخرى طبقتان.

إحدهما طبقة واحدة مخفاة، بينما تستخدم الأخرى طبقتين مخفاتيْن.

وتتلخص مرحلة تدريب الشبكة العصبية في عملية إيجاد مجموعة مناسبة من الأوزان لجميع الوصلات بين العقد بحيث يمكن توليد المخرجات المطلوبة المقابلة لمدخلات معينة. ولم يكن هذا النوع من الشبكات العصبية ذات الإدراك المتعدد الطبقات مستخدماً من قبل وذلك بسبب نقص خوارزميات التدريب الفعالة المناسبة لها. ولكن هذه الحالة قد تغيرت مؤخراً كنتيجة مباشرة لتطوير إجراء متدرج تكراري يعرف باسم خوارزمية الانتشار الرجوعي Back-propagation algorithm. وطبقاً لهذه الخوارزمية، التي سيتم شرح خطوطها الرئيسة بعد قليل، فإن التعلم - أي الحصول على أوزان مناسبة لكل نموذج أو شكل من أشكال مجموعة التدريب - يتم بحساب الفروق بين المخرجات المطلوبة والمخرجات الفعلية، ثم تغذية إشارة الخطأ هذه - رجوعاً - إلى المدخلات، لكل مستوى على حدة، وذلك لتغيير أوزان الوصلات المعينة بطريقة تضمن تعديل هذه الأوزان بدرجة تناسب مسؤولية كل منها عن هذا الخطأ المعين. وفيما يلي تجريد للخطوات الرئيسة لهذه الخوارزمية [١٩]:

الخطوة (١): أعط قيماً استهلالية لجميع الأوزان W_{ij} والمقابلة للوصلات بين كل عقدة z وعقدة i في الطبقة التي تحتها مباشرة.

الخطوة (٢): قدم إحدى القيم المدخلة من الفئة m وحدد قيم المخرجات المطلوبة مقابلها .

الخطوة (٣): احسب قيم المخرجات الفعلية لجميع العقد باستخدام قيم الأوزان W_{ij} الحالية، ويكون المخرج من العقدة j والمسمى y_j دالة غير خطية في جميع المدخلات الحالية كما يلي :

$$y_j = \frac{1}{1 + e^{- (\sum_i y_i W_{ij} + \theta_j)}} \quad (10)$$

حيث θ هي قيمة الانحياز Bias وهذه الدالة غير الخطية المستخدمة هنا تسمى الدالة الأسية Sigmoid function .

الخطوة (٤): أوجد حد الخطأ δ_j لجميع العقد . فإذا مثلنا المخرج المطلوب بالقيمة d_j والمخرج الفعلي بالقيمة y_j وذلك عند عقدة معينة، فإن حد الخطأ δ_j عند عقدة مخرجات يكون

$$\delta_j = (d_j - y_j) y_j (1 - y_j) \quad (11)$$

بينما يكون حد الخطأ δ_j عند عقدة في طبقة مخفاة

$$\delta_j = y_j (1 - y_j) \sum_k \delta_k w_{jk} \quad (12)$$

حيث k هي جميع عقد الطبقة الأعلى من الطبقة j .

الخطوة (٥): اضبط قيم الأوزان كما يلي

$$W_{ij}(n+1) = W_{ij}(n) + \alpha \delta_j y_i + \epsilon (W_{ij}(n) - W_{ij}(n-1)) \quad (13)$$

حيث إن :

$(n+1), n, (n-1)$ هي أدلة على الأوزان السابقة والحالية على الترتيب، α هي معدل التعلم ويأخذ قيمة بين الصفر والواحد الصحيح، ϵ هو ثابت قيمته بين الصفر والواحد الصحيح وهو الذي يحدد تأثير التغييرات السابقة للأوزان على الاتجاه الحالي لحركة الوزن في حيز الأوزان. ويتيح هذا نوعاً من القوة الدافعة التي تسهم في استبعاد التغييرات ذات التردد العالي في قيمة الخطأ.

الخطوة (٦): قدم قيمة مدخلة أخرى ثم ارجع إلى الخطوة رقم (٢). ويتوالى تقديم جميع المدخلات الخاصة بالتدريب (التعليم) بطريقة دورية حتى تتقارب قيم الأوزان المستخدمة.

٧. التجارب والنتائج

انقسم العمل التجريبي في هذا البحث إلى ثلاثة أجزاء. تناول الجزء الأول التجارب التي تم فيها استخدام المصنفات الإحصائية التقليدية، بينما تناول الجزء الثاني من العمل التجريبي استخدام الشبكات العصبية ذات الانتشار الرجوعي. أما الجزء الثالث فتمت فيه مقارنة النتائج التي تم التوصل إليها باستخدام كل من الطريقتين سالفتي الذكر.

لقد تم استخدام أربع مجموعات مختلفة من أشكال حروف العربية، وكل مجموعة من هذه بها ٧٨ حرفاً مختلفاً. وقد تم تمثيل كل حرف في ملف البيانات باستخدام مجموعات من الرقم ١ وفراغات وذلك بطريقة مناسبة. وأمكن تقليل الأشكال المميزة للحروف في كل مجموعة، وذلك باستخلاص النقاط المستخدمة مع الحروف قبل مرحلة التعرف، وبهذه الطريقة فإن الحروف المتشابهة والتي كانت تميزها النقاط أصبحت نموذجاً واحداً. فمثلاً الحروف المختلفة مثل (ب، ت، ث) تم اختصارها في نموذج واحد هو (ب). وبعد تمام التعرف على هذا النموذج يمكن استخدام النقاط لتمييز الحرف الأصلي. وعليه فقد أمكن تخفيض عدد الحروف في كل مجموعة من ٧٨ حرفاً إلى ٤٤ نموذجاً فقط.

وبتطبيق تحويل " والش " على الإسقاطين الأفقي والرأسي ، وذلك باستخدام بيانات الرسم البياني النسيجي الأفقي والرأسي لكل نموذج من النماذج ، تم الحصول على الملامح الخاصة بنماذج كل مجموعة من المجموعات الأربع وخزنت في ملف خاص . وبالتالي فقد حصلنا على أربعة ملفات ، تم تقسيمها إلى قسمين : قسم للتدريب (التعليم) وقسم للاختبار والتجربة . أما العزوم المركزية فقد تم التجهيز لاستخدامها بالطريقة المذكورة أعلاه ، عدا أن كل نموذج من نماذج حروف العربية يحتاج هنا إلى سبعة ملامح (بدلاً من عشرين ملمحاً بطريقة تحويل " والش ") .

١ ، ٧ . المصنفات الإحصائية التقليدية

لقد تم في هذا البحث اختيار أداء المصنف الإحصائي المعروف بالمسافة الصغرى (باستخدام القيمة المطلقة للفرق) وذلك للمقارنة مع أداء الشبكة العصبية ذات الانتشار الرجوعي . وسنقدم هنا شرحاً مختصراً لهذا المصنف الإحصائي .
افرض أن لدينا n من الفئات المختلفة $(C_1, C_2, C_3, C_4, \dots, C_n)$ ، كل منها له عدد m من الملامح التي يمكن استخدامها ، أي أن

$$C_i = (f_{i1}, f_{i2}, \dots, f_{im})$$

حيث C_i هو الفئة رقم i . فإذا أردنا تصنيف عينة مجهولة ، فإنه يلزم حساب مجموع القيم المطلوبة للفرق بين قيمة كل ملمح من ملامح العينة x وبين قيمة نفس الملمح للفئة i ، ويعبر عن ذلك رياضياً بالصيغة :

$$SDA(x, i) = \sum_{j=1}^m |X_j - M_{ij}|$$

حيث M_{ij} تدل على الملمح j للنموذج i . ويتم حساب هذه المسافة بين كل من العينة المعطاة x (المطلوب التعرف عليها) وبين جميع النماذج الموجودة في قاعدة البيانات . وبعد ذلك يمكن إيجاد المسافة الصغرى باستخدام المعادلة

$$C_{min} = \text{Min} [SDA(x, i)], \quad i = 1, 2, 3, \dots, n$$

وتستخدم هذه القيمة C_{min} في نسبة العينة المجهولة إلى فئة النموذج المعين والذي بينه وبين العينة المسافة الصغرى كما سبق .

يبين الجدولان رقماً (١, ٢) ملخصاً للنتائج التجريبية التي تم الحصول عليها باستخدام المصنفات الإحصائية التقليدية، حيث يلخص الجدول رقم (١) النتائج التي تم الحصول عليها مع استخدام مجموعة واحدة من الحروف بغرض عمل النماذج المستخدمة في التمييز، أما الجدول رقم (٢) فيلخص النتائج المشابهة للمذكورة آنفاً مع استخدام مجموعتين من الحروف بغرض عمل النماذج المستخدمة في التعرف على الحروف .

جدول رقم ١ . مختصر التجارب التي أجريت باستخدام المصنف التقليدي واستخدام مجموعة واحدة من الحروف لعمل النماذج .

نسبة التعرف % للمجموعة				
4	3	2	نوع المجموعة	الملامح المستخدمة
87.18	93.59	96.15	حروف	والش
88.64	95.45	100	نماذج الحروف	والش
42.31	41.03	60.26	حروف	عزوم
43.18	54.55	59.09	نماذج الحروف	عزوم

جدول رقم ٢ . مختصر التجارب التي أجريت باستخدام المصنف التقليدي واستخدام مجموعتين من الحروف لعمل النماذج .

نسبة التعرف % للمجموعة			
3	2	نوع المجموعة	الملامح المستخدمة
98.72	98.72	حروف	والش
100	100	نماذج الحروف	والش
70.26	67.69	حروف	عزوم
79.55	81.82	نماذج الحروف	عزوم

استخدمت ملامح كل من طريقتي " والش " و " العزوم " في كل من مرحلتي التدريب والاختبار . وفي إحدى التجارب استخدمت مجموعة بيانات الحروف رقم (١) في التدريب ثم استخدمت المجموعات أرقام (٢، ٣، ٤) للاختبار . ويوضح الجدول رقم (١) نتائج الاختبارات التجريبية التي تمت باستخدام كل من مجموعات الحروف ومجموعات نماذج الحروف . وقد تم الحصول على نماذج الحروف العربية في هذا البحث بتجريد الحروف المتشابهة من النقاط المميزة لها والإبقاء على الجزء الأساسي المشترك بين هذه الحروف المتشابهة . أما النقاط فقد تم استخلاصها في مرحلة معالجة سابقة باستخدام الحدود المحيطية Contours . ويتم عند التجربة التعرف على حرف العربية المجهول باستخدام رقم النموذج المشابه ثم عدد النقاط الموجودة مع الحرف وموقعها . أما النتائج الملخصة في الجدول رقم (٢) فقد تم الحصول عليها من مجموعة تجارب أخرى قصد بها تحسين معدلات التعرف ، حيث تم التدريب باستخدام اثنتين من مجموعات بيانات الحروف ، وهما على وجه التحديد المجموعتان رقما (١ ، ٤) ، واستخدمت المجموعتان رقما (٢ ، ٣) في الاختبارات واتضح من ذلك - وبالمقارنة مع الجدول رقم (١) - ارتفاع معدلات التمييز وتحسنها في حالة استخدام مجموعتين للتدريب بدلاً من مجموعة واحدة .

٧,٢ . مصنف الشبكة العصبية ذات الانتشار الرجوعي :

لكي تكون مصنفات الشبكة العصبية بديلاً مقبولاً للمصنفات الإحصائية التقليدية ، ينبغي للشبكة العصبية أن يتفوق أداؤها على - أو في الأقل يتساوى مع - أداء المصنفات الإحصائية التقليدية .

ولقد شرحنا نموذج الشبكة العصبية ذات الانتشار الرجوعي في الفصل رقم (٦) من هذا البحث ، وتم استخدام تراكيب بنيوية مختلفة لهذا النوع من الشبكات العصبية في تجارب مختلفة . فمثلاً عندما استخدمنا ملامح تحويل " والش " كان عدد عقد طبقة المدخلات هو (٢٠) ، بينما تم اختيار أعداد مختلفة من عقد الطبقة المخفاة (٢٠ ، ٤٠ ، ٦٠ ، ٨٠ ، ١٠٠ عقدة) ، وكان عدد عقد المخرجات ٧٨ عقدة عند استخدام كل حروف المجموعة و ٤٤ فقط عند استخدام مجموعة النماذج . وفي هذا البحث تم استخدام

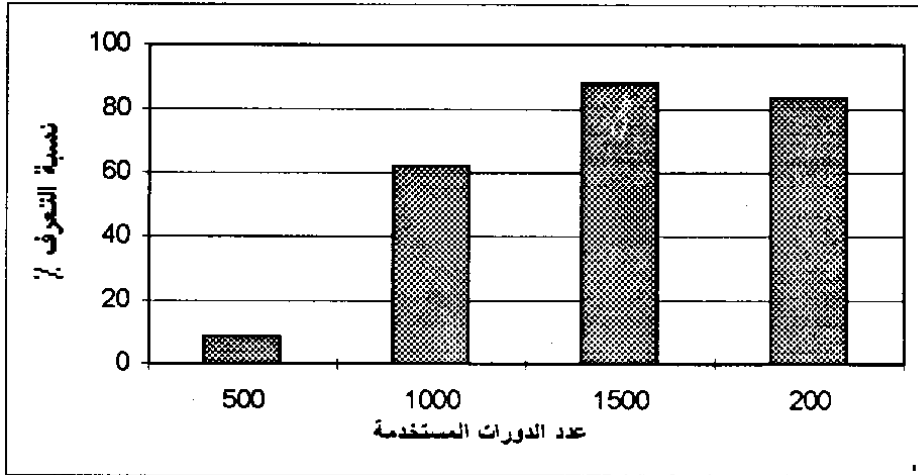
نموذج الشبكة العصبية ذات الانتشار الرجوعي كمصنف، بديلا للمصنفات الإحصائية التقليدية، كما تم عمل مقارنة بين أداء كل منهما .
وقد استخدمت في هذه الحالة مجموعة واحدة من حروف العربية لتدريب الشبكة العصبية وادخرت المجموعات الثلاث الباقية للاختبار، ويلخص الجدول رقم (٣) النتائج التجريبية التي تم التوصل إليها باستخدام ملامح " والش " .

جدول رقم ٣ . مختصر التجارب التي أجريت باستخدام ملامح والش واستخدام مجموعة من الحروف لتدريب الشبكة العصبية ذات الطبقتين المخفتين واستخدام أعداد مختلفة من العقد المخفاة .

نسبة التعرف % مع اختلاف عدد الدورات								رقم المجموعة المستخدمة في الفحص	عدد العقد المخفاة في كل من الطبقتين	
4000	3500	3000	2500	2000	1500	1000	500		2	1
--	--	86.4	22.7	36.4	43.1	9.1	0	2	60	40
--	--	77.3	20.5	25.0	27.3	6.8	0	3	60	40
--	--	72.7	18.2	18.2	25.0	6.8	0	4	60	40
--	--	--	--	--	13	2.3	0	2	100	40
75.0	77.3	40.9	4.6	43.2	9.1	2.3	0	3	100	40
65.9	70.5	45.5	2.3	36.4	9.1	2.3	0	4	100	40
--	--	93.2	86.4	59.1	27.3	2.3	0	2	60	80
--	--	84.1	81.8	54.6	20.5	2.3	0	3	60	80
--	--	65.9	68.2	52.3	18.2	2.3	0	4	60	80
--	88.6	9.1	47.7	9.1	0	0	0	2	100	80
--	80.0	6.82	40.9	6.8	0	0	0	3	100	80
--	70.5	9.09	31.8	6.8	0	0	0	4	100	80

وتتضح من الجدول رقم (٣) والشكل رقم (٤) معدلات التعرف على مجموعة الحروف رقم (٢) التي سجلت باستخدام الشبكة العصبية، ذات ٦٠ عقدة في الطبقة المخفاة . ويوضح الشكل أن معدل التعرف كان ٦، ٨٨٪ عند ١٥٠٠ دورة ثم انخفضت قيمته إلى ٠٩، ٨٤٪ عند ٢٠٠٠ دورة، مما يشير إلى حدوث مطابقة زائدة Overfitting للنموذج (الأوزان) مع بيانات الحروف عند ٢٠٠٠ دورة . ويؤكد هذا بالتالي أن الأوزان

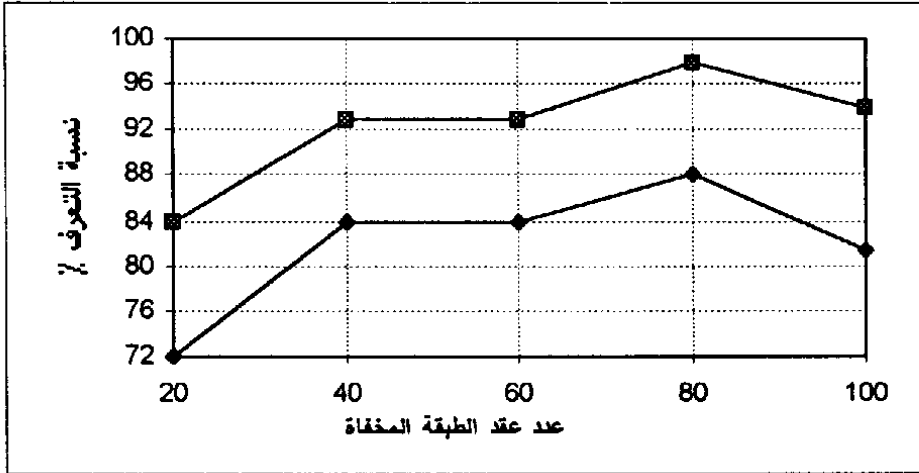
التي حصلنا عليها عند ١٥٠٠ دورة هي الأفضل لهذا التدريب. أما عند اختلاف مجموعات بيانات الحروف المستخدمة فإن الأوزان عند ١٥٠٠ دورة ليست بالأفضل. وهذا يجعل استخدام مجموعات حروف متعددة عند تدريب الشبكات العصبية أمراً ضرورياً.



شكل رقم ٤. معدل التعرف على حروف المجموعة رقم ٢ باستخدام ملامح «والش» و ٦٠ عقدة في الطبقة المخفأة.

ومن أجل تحليل تأثير عدد العقد في الطبقة المخفأة على عملية التعرف، تمت تجربة عدد من الشبكات العصبية ذات أعداد مختلفة من العقد في طبقاتها المخفأة. ويوضح شكل رقم (٥) العلاقة بين معدل التعرف وعدد العقد في الطبقة المخفأة، ويظهر جلياً أن معدل التعرف يزيد بزيادة عدد العقد من ٢٠ حتى ٨٠ عقدة ثم يبدأ في الانخفاض مرة أخرى.

ولقد تم استخدام اثنتين من مجموعات نماذج الحروف، وهما بالتحديد المجموعتان رقم (١، ٤) لتدريب الشبكة العصبية المستخدمة في التجربة بينما تم استخدام المجموعتين الأخرين رقم (٢، ٣) للاختبار. واتضح من ذلك أن زيادة عدد المجموعات المستخدمة للتدريب إلى اثنتين قد حسن معدلات التعرف على كل مجموعات حروف



شكل رقم ٥ . معدل التعرف باستخدام مجموعة واحدة ثم مجموعتين من الحروف :
 ◆ مجموعة واحدة □ مجموعتان

الاختبار في الوقت الذي تم فيه الحصول على الأوزان المطلوبة للتعرف باستخدام عدد الدورات نفسه أو أقل منه، ويوضح الجدول رقم (٤) معدلات التعرف التي تم التوصل إليها نتيجة لإجراء الاختبارات.

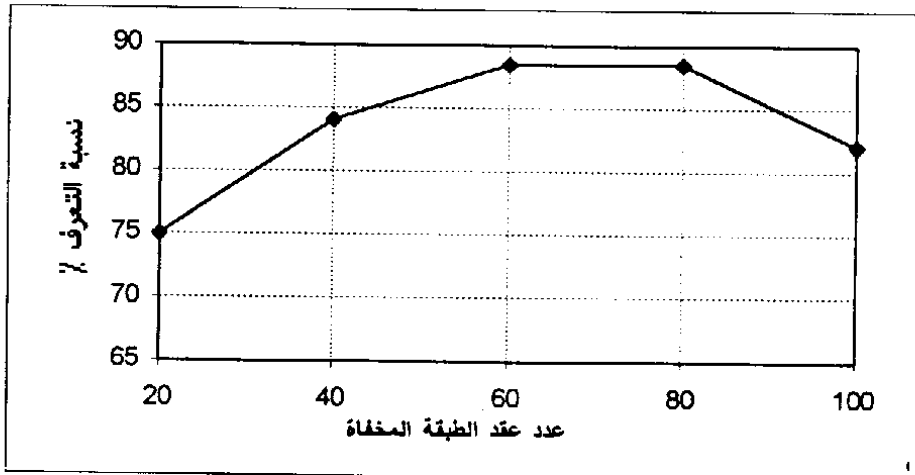
وكما ذكرنا سابقاً، تم استخدام شبكات عصبية بأعداد مختلفة من العقد في الطبقة المخفأة وذلك بغرض دراسة تأثير هذا العدد على معدلات التعرف. ويوضح شكل رقم (٦) النتائج التي تم الحصول عليها في هذا الصدد. ولقد اتضح من هذه التجارب أن زيادة عدد الطبقات المخفأة إلى طبقتين قد أدى إلى تحسين معدلات التعرف على كل نماذج الاختبارات إلا أن الأوزان اللازمة للتعرف قد استغرق الحصول عليها عدداً أكبر من الدورات، مما استلزم زيادة مطردة في كل من متطلبات التخزين والحسابات الآلية.

٧,٣ . مقارنة بين استخدام كل من الشبكات العصبية ذات الانتشار الرجوعي والمصنفات الإحصائية التقليدية:

هناك عوامل عديدة تتسبب في تعقيد عملية المقارنة بين استخدام الشبكات العصبية ذات الانتشار الرجوعي واستخدام المصنفات الإحصائية التقليدية. فمن المعروف أن

جدول ٤ . مختصر التجارب التي أجريت باستخدام ملامح والش واستخدام مجموعتين من الحروف لتدريب الشبكة العصبية ذات الطبقة المخفاة واستخدام أعداد مختلفة من العقد المخفاة.

نسبة التعرف % منسوبة إلى عدد الدورات				رقم المجموعة المستخدمة في الفحص	عدد العقد في الطبقة المخفاة
2000	1500	1000	500		
84.09	50.00	45.45	13.64	2	20
88.64	54.55	45.45	11.36	3	20
--	93.18	50.00	9.09	2	40
--	88.64	52.27	9.09	3	40
--	93.18	47.73	11.36	2	60
--	88.64	36.36	11.36	3	60
--	97.73	70.45	6.82	2	80
--	88.64	70.45	6.82	3	80
93.82	70.55	38.64	0	2	100
95.45	70.55	38.64	0	3	100



شكل رقم ٦ . العلاقة بين معدل التعرف وعدد العقد في الطبقة المخفاة

المصنفات الخاصة بالشبكات العصبية ذات الانتشار الرجوعي تعتمد على عمارة هذه الشبكات . وبتعبير آخر فإنها - أي المصنفات - تعتمد على عدد الطبقات المخفأة، وعدد العقد في كل من هذه الطبقات المخفأة، بالإضافة إلى اعتمادها على عوامل أخرى . وعلاوة على ذلك فإن الشبكات العصبية تعاني من مشكلات عدة من بينها التقارب Convergence والتعميم Generalization وكذا تعريف التقارب في حد ذاته . ومن المعروف أن التقارب المطلق لا يؤدي بالضرورة إلى أفضل تعميم ممكن، كما أن تأثير إضافة وحدات مخفأة على التعميم لم يمكننا سبر غوره كاملاً بعد . ويمكننا - من النتائج التي سقناها في هذا البحث - أن ندرك قابلية أفضل أداء، في حقل التعرف على حروف العربية باستخدام المصنفات الخاصة بالشبكات العصبية ذات الانتشار الرجوعي، للمقارنة مع أداء المصنفات الإحصائية التقليدية . وعلى أية حال، يجب أن نأخذ في اعتبارنا أنه بينما يسهل استخدام المصنفات الإحصائية التقليدية نظراً لكونها مباشرة التطبيق، فإن البنية التركيبية المثلى للشبكة العصبية ذات الانتشار الرجوعي لا يمكن الحصول عليها إلا بمحاولات عديدة، وعندئذ فقط يمكننا أن نقول بإمكانية الحصول على تحسن كبير في أداء هذه الشبكات العصبية بالمقارنة مع المصنفات الإحصائية التقليدية .

٨ . خاتمة البحث

لقد تم في هذا البحث تطوير نظام للتعرف على حروف العربية باستخدام شبكة عصبية ذات انتشار رجوعي . ومن المعلوم أن التعرف على حروف العربية أصعب من التعرف على حروف اللغات اللاتينية لاختلاف حروف العربية عن غيرها من حروف اللغات الأخرى في البنية التركيبية وفي طريقة توصيلها وقد اتبعت في هذا البحث منهجيتان هما تحويل " والش " وطريقة " العزوم " . وقد أدى استخدام الطريقة الأولى - أي تحويل " والش " - إلى تعرف أفضل وأسرع من الطريقة الثانية، غير أن استخدام الطريقة الثانية - طريقة العزوم - يحتاج إلى مساحة تخزين أقل .

استخدمت عمارات مختلفة للشبكة العصبية ذات الانتشار الرجوعي في تجارب مختلفة، وذلك لدراسة تأثير عدد العقد في الطبقة المخفأة على أداء الشبكة عند التعرف

على حروف العربية، كما تمت دراسة تأثير عدد المجموعات المستخدمة في تدريب الشبكة على عملية التعرف على الحروف وقد تمت أيضاً دراسة تأثير عدد الطبقات المخفاة على عملية التعرف. ولكي تكون الشبكة العصبية بديلاً مقبولاً للمصنفات الإحصائية التقليدية، يجب أن يتفوق أداء الأولى على الثانية، أو أن يكون مساوياً له في الأقل. ولغرض المقارنة بين الطريقتين، تم اختيار المسافة الصغرى باستخدام القيمة المطلقة للفروق كمصنف إحصائي تقليدي، وتبين استوائه في الأداء مع الشبكة العصبية ذات الانتشار الرجوعي. ولكن ينبغي أن ندرك أنه في الوقت الذي يتم فيه تطبيق المصنفات الإحصائية التقليدية بصورة مباشرة فإن التركيب الأمثل للشبكة العصبية ذات الانتشار الرجوعي لا يمكن التوصل إليه إلا بعد بحث مستفيض، وإن كانت الأخيرة تتمتع بإمكانات واسعة لتحسين الأداء بالإضافة إلى كونها قوية الاحتمال للتشويش.

ولقد تبين أن زيادة عدد مجموعات التدريب قد أدى إلى تحسين معدلات التعرف في كلتا الطريقتين، أي المصنفات التقليدية والشبكات العصبية.

تم تطوير هذا النظام باستخدام لغة سي (C) تحت نظامي تشغيل دوس (DOS) ويونيكس (UNIX). والنظام الأخير مركب على محطة عمل "صن" (Sun Workstation) المحتوية على أربعة معالجات. وقد كان تدريب الشبكة العصبية باستخدام محطة عمل "صن" أسرع بحوالي ٣٠ مرة منه باستخدام نظام تشغيل "دوس" المركب على حاسب شخصي بمعالج ٨٠٣٨٦ هذا بالإضافة إلى أن نظام "يونيكس" يتيح تشغيل عدة برامج في آن واحد وهذا غير ممكن مع نظام "دوس".

شكر وتقدير: يود الباحثان أن يعربا عن شكرهما وتقديرهما للمحكمين، اللذين أدى نقد كل منهما البناء إلى زيادة وضوح عرض نتائج هذا البحث.

٩. المراجع

- [١] Mantas, J. "An Overview of Character Recognition Methodologies." *Pattern Recognition*, 19, No. 6 (1986), 425-430.
- [٢] Govindan, V.K. and Shivaprasad, A.P. "Character Recognition-A Review." *Pattern Recognition*, 23, No.7 (1990), 671-683.

- Yousefi,H.S. and Upda,S.S. "Recognition of Arabic Characters." *IEEE Trans. on PAMI*, 14, No. 8 (Aug. 1992),853-857. [٣]
- El-Sheikh,T. and Guindi, R."Computer Recognition of Arabic Cursive Scripts." *Pattern Recognition*, 21,No.4 (1988), 293-302. [٤]
- El-Khaly,F. and Sid-Ahmed,M."Machine Recognition of Optically Captured Machine Printed Arabic Text." *Pattern Recognition*,23,No. 11 (1990), 1207-1214. [٥]
- El-Dabi, S. S.; Ramsis, R. R. and Kamel, A. "Arabic Character Recognition System: A Statistical Approach for Recognizing Cursive Typewritten Text." *Pattern Recognition*, 23,No. 5 (1990), 485-495. [٦]
- Almuallim, H. and Yamaguchi, S. "A Method of Recognition of Arabic Cursive Handwriting." *IEEE Trans. Pattern. Anal. Machine Intell.*, PAMI-9, No.5 (Sept. 1987), 715-722. [٧]
- El-Sheikh, T.S. and El-Taweel, S.G."Real-time Arabic Handwritten Character Recognition." *Pattern Recognition*, 23,No.12 (1990), 1323-1332. [٨]
- El-Sheikh,T.S. and El-Taweel,S.G."Segmentation of Handwritten Words." *Proceedings of the 12th National Computer Conference*, Riyadh, Saudi Arabia (Oct., 1990), 389-402. [٩]
- El-Sheikh,T.S. and El-Taweel, S.G. "Recognition of Typewritten Arabic Characters in Different Fonts." *Proceedings of the IEE Colloquium on Character Recognition and Applications*, London, England (Oct., 1989),9/1-9/5. [١٠]
- El-Gowely, K.; Dessouki, O. and Nazif, A."Multi-phase Recognition of Multifont Photoscript Arabic Text." *Proc.IEEE* (1990), 700-702. [١١]
- Amin, A. and Al-Fedaghi, S. "Machine Recognition of Printed Arabic Text Utilizing Natural Language Morphology." *Int. Man-Machine Studies*,Vol. 35, No. 6 (1991), 769-788. [١٢]
- Amin, A. and Masini, G."Machine Recognition of Multifont Printed Arabic Texts." *IEEE Proc.* (1986), 392-395. [١٣]

- Amin, A. "Machine Recognition of Handwritten Arabic Words." *IEEE Proc* [١٤]
(1982), 34-36.
- Goraine, H.; Usher, M. and El-Emami, S. " Off-Line Arabic Character Recognition. " *IEEE Computer*, 25, No. 7 (July 1992), 71- 74. [١٥]
- Gonzalez, R.C. and Wintz, P. "*Digital Image Processing*." Addison-Welsey, [١٦]
1977.
- Schalkoff, R. J. *Digital Image Processing and Computer Vision*. John Wiley [١٧]
& Sons, Inc., New York: 1989 , 304-307.
- Lippmann, R. P. "An Introduction to Computing with Neural Nets." *IEEE* [١٨]
ASSP Magazine , Vol. 4, No.2 (April 1987), 4-22.
- Hush, Don R. and Horne, Bill G. "Progress in Supervised Neural Networks, [١٩]
What's New Since Lippmann?" *IEEE Signal Processing Magazine* (Jan. 1993),
8-39.

Recognition of Typewritten Arabic Characters Using Backpropagation Neural Networks

Asim Abdul-Fattah Nabawi *; Sabri Abdullah Mahmoud **

** Computer Engineering Dept. CCIS, King Saud University, Riyadh*

*** Al-Manarain Est. for Technical Applications, P.O. Box 53531, Riyadh 11593*

Abstract. Research efforts in the field of Arabic character recognition using optical techniques have intensified in the recent past. This state of affairs prevail due to several reasons, among which is the increasing interest in computer systems combined with their affordable availability to a broader class of users, not to mention the technical advances in OCR applications in languages other than Arabic.

This paper addresses the optical recognition of Arabic characters using the well known backpropagation neural network. Features of the Arabic characters based on Walsh transform and the moments method were extracted. Both classical and neural network classifiers were used in both the training and recognition phases. Comparison of the two classification techniques and the two types of extracted features was carried out.

The results of this research work have shown that the use of the Walsh-transform-based features yielded higher recognition rates in addition to faster processing times compared to those of the moments-based technique. The latter technique, however, requires less memory. The backpropagation technique was found to be more tolerant to noise than the classical classifiers. The former, however, was found to require several trials to adapt its architecture to the needs of the particular application. The experimental results have ascertained the effect of using larger sets of characters on improving the recognition rates in all instances.