



جامعة السودان للعلوم والتكنولوجيا

كلية الدراسات العليا

التعرف على الأسماء العربية المكتوبة باستخدام الطريقة الشمولية

Arabic Handwritten Names Recognition by Using Holistic Approach

إعداد الطالبة:

وفاء علي محمد عبد الرحمن

إشراف:

د. محمد الحافظ مصطفى موسى

بحث مقدم كأحد متطلبات الحصول على درجة الماجستير في علوم الحاسوب

سبتمبر 2008

الحمد لله

الحمد لله رب العرش الكريم خالق الملائكة وجاعل النبيين بشراً من طين

الحمد لله الذي جعل الليل و النهار آيتين متعاقبتين

الحمد لله خالق السموات والأرض

الحمد لله الذي أحسن كل شيء صنعا

الحمد لله الذي جعل القرآن عربيا غير ذي عوج

الحمد لله الذي يستحق الحمد.

الآية

قَالَ تَعَالَى:

﴿وَلَقَدْ نَعَلِمَ أَنَّهُمْ يَقُولُونَ إِنَّمَا يُعَلِّمُهُ بَشَرٌ لِّسَانٌ الَّذِي يُلْحِدُونَ إِلَيْهِ أَعْجَبِي وَهَذَا لِسَانٌ

عَرَبِيٌّ مُّبِينٌ ﴿١٠٣﴾

النحل: ١٠٣

الإهداء

إلى من تعهدني بالرعاية وكان دعماً لي في جميع مراحل حياتي
والذي العزيز

إلى من كانت دعواتها تذلل الصعاب
والذي الحبيبة

إلى إخواني و أخواتي
إلى من يحتلون مساحة من قلبي
اليكم اهدي
وفاء

الشكر و العرفان

الشكر أوله وآخره لله الواحد المتعال الذي انعم علينا ومنحنا العقل.
ومن ثم الشكر أجله للدكتور محمد الحافظ مصطفى الذي ما بخل بزمن أو جهد أو معلومة وكان نعم المشرف في كل الخطوات.
والشكر لمجموعة التعرف على الأنماط البحثية بجامعة السودان للعلوم والتكنولوجيا واطح بالشكر الأخت بدور علمي التي كانت لها مواقف مقدرة في تكوين قاعدة البيانات الخاصة بالأسماء.
والشكر لأساتذة و طلاب كلية علوم الحاسوب و تقانة المعلومات بجامعة السودان للعلوم والتكنولوجيا لما قدموه من دعم.
والشكر لكل من منحني دعماً معنوياً للمواصلة و إخراج هذا البحث.
وأخرا الشكر لكل من مد يد العون و ساهم في إخراج هذا العمل المتواضع.

المستخلص

بناء حواسيب تستغني عن أدوات الإدخال محدودة الفضاء (مثل لوحة المفاتيح) وامتلاكها لمقدرة السمع و القراءة ظل من مجالات البحث النشطة في علوم الحاسوب لأكثر من أربعة عقود، قدم فيها الباحثون عدد مقدر من الطرق و الخوارزميات لحوسبة السمع و القراءة ضمن ما يعرف بالتعرف على الأنماط (Pattern Recognition) في علوم الحاسوب. ومن بين هذه الطرق الطريقة الشمولية، التي أثبتت كفاءتها في التعرف السريع (سمعاً أو قراءة) على عدد محدد من الكلمات. حديثاً زاد الاهتمام بهذه الطريقة بعد الزيادة الكبيرة في سرعة المعالجة الحاسوبية و التقدم في المعالجة المتوازية. هذه الدراسة تقدم تجارب إدراك ناجحة للشبكات العصبية الاحتمالية في التعرف شمولياً على الأسماء العربية الأكثر شيوعاً، حيث تم تصميم شبكة احتمالية و تمت تجربتها على الخمسة عشر اسم الأكثر شيوعاً بحسب مجموعة بيانات جامعة السودان للاسماء (SUST-ARG names) شملت التجارب اختيار عتبة للرفض حتى لا يتم اعتماد التعرف على الاسماء بدرجة احتمالية ضعيفة، تعد نتيجة التجارب جيدة جدا حيث كان معدل الخطأ اقل من 1% مع معدل رفض 34%.

ABSTRACT

Designing computerized systems which possess reading and hearing faculties is an active research area for more than four decades. Many methods and algorithms have been suggested by researchers for this purpose as part of pattern recognition research. Recently, more research work has been devoted to the holistic approach, as a result of the rapid growth of computation speed and memory. In the holist approach the recognition system recognizes a complete word as one object with going through the long and erroneous letter segmentation process. In this study a probabilistic neural network has been designed to recognize the popular Arabic names holistically. SUST ARG names data set has been used to test the network performance. Selecting an appropriate threshold for rejection has been investigated. With a threshold value which rejects 34% of the testing data the network has less than 1% error rate.