محمود ; المجلة العربية للعلوم و نشر الأبحاث، المجلد الأول - العدد (3): 21- أبريل 2016؛ ص. 1-11 ؛ بحث رقم:D14216 .

تصميم شبكة عصبية اصطناعية للتحكم الصوتي في منظومة تحكم بضغط أنابيب توزيع المنتجات النفطية

دراسة بحثية

الدكتورة/ داليا محمود آدم محمود

جامعة النيلين ، كلية الهندسة ، الخرطوم ، السودان

البريد الإلكتروني : d.daliamahmoud@yahoo.com

الملخص :

تقدم هذه الورقة دراسة اجريت لتصميم شبكة عصبية اصطناعية للتحكم الصوتي في منظومة تحكم بالضغط حيث تمت الاستفادة من قوة الشبكات العصبية الاصطناعية وهي إحدى أفرع علم الذكاء الاصطناعي AI في بناء متعرف صوتي لاستخدامه داخل منظومة تحكم في ضغط أنابيب توزيع المنتجات النفطية. قدمت الورقة اسلوباً فعالاً في تصميم الشبكة العصبية للحصول على أفضل اداء ممكن من التصميم .

الكلمات المفتاحية: التعرف على الكلام، بارامترات LPC ، خوارزمية Kmean .

1. مقدمة

أحد أهداف الشبكات العصبية الاصطناعية هو معالجة المعلومة بأسلوب مشابه لذلك المتبع في الأنظمة البيولوجية الحية، وحقيقة فأننا نستخدم الشبكات العصبية عندما نحتاج إلى قدرات الدماغ ومثالية الآلة.
التصميم ومن ثم التدريب الجيدان للشبكة العصبية قد يؤهلها لأداء عمليات اتخاذ القرار عند مواجهتها لمشاكل أو ظروف معينة خارج نطاق بيانات التعليم التي دربت عليها، وهذا يكسب الشبكات العصبية وثوقية عالية تماماً كما نثق في قرارات الخبراء من البشر.

هنالك أربع أهداف أساسية في مجال البحث في الشبكات العصبية الاصطناعية [5] وهي:

1. هيكلة الشبكة (Network structure).
2. هدف التصميم (Design objective).
3. تحسين الاداء (Optimization).
4. قوة الشبكة عند التعامل مع حالات جديدة وتعرف هذه بعمومية الشبكة (Generalization).

وقد ركز هذا العمل على الهدفين الاول والرابع لما لهما من أثر في تحسين أداء الشبكة العصبية.

هيكلة الشبكة

المشكلة الاساسية عند تصميم شبكة عصبية اصطناعية هو اختيار البنية المناسبة Neural Network Architecture للتطبيق المراد ويشمل ذلك ما يلي :

1. اختيار نوع الشبكة الملائم للتطبيق.
2. اختيار عدد الطبقات.
3. اختيار عدد العقد المخفية بكل طبقة.
4. اختيار خوارزمية تفعيل الخرج (دالة الانتقال) لكل طبقة.

بالنسبة لتطبيقات التعرف على الكلام فان العديد من الدراسات والابحاث قد اجريت لتحديد أي أنواع الشبكات أنسب لاداء عملية التعرف على الكلام وقد أثبتت تلك الدراسات أن الشبكات العصبية ذات التغذية المرتدة (Feed Forward Back Propagation Networks) تعطي أفضل نسب اداء عند التعرف على الكلام[5]. أما بالنسبة لاختيار عدد الطبقات وعدد العقد وخوارزمية التفعيل فان ذلك يتحدد بناء على التطبيق المراد استخدام الشبكة العصبية فيه، ولا توجد قاعدة معينة للاختيار. وفي هذا البحث تم الاختيار بناء على طريقة التجربة والخطأ (Trail and Error) وهي طريقة موضحة في الدراسة [6] حيث تم اختيار عدد ابتدائي من الطبقات والعقد وتم اختيار احدى خوارزميات التفعيل ورصدت قيمة الخطأ في الاداء، ثم تم تغيير دالة التفعيل لنفس الشبكة ورصدت قيمة الخطأ في الاداء مرة اخرى، وتكرر العملية مره اخرى لدالة تفعيل جديدة. وفي النهاية تم اختيار دالة التفعيل التي اعطت أقل قيمة خطأ في الاداء. بعد ذلك تم تحديد عدد جديد من العقد ذو قيمة أقل من العدد الابتدائي ثم رصدت قيمة الخطأ في الاداء، وتم تحديد عدد أكبر للعقد من العدد الابتدائي وتم رصد قيمة الخطأ في الاداء والغرض من هذا تحديد هل يحتاج التطبيق الى عدد أقل أو أكبر من العقد، فاذا ظلت قيمة الخطأ في الاداء عالية فأن هذا يتطلب زيادة عدد الطبقات وتكرار عملية اختيار العقد بكل طبقة. وفي النهاية تختار الشبكة العصبية التي أعطت أقل قيمة خطأ في الاداء.

عمومية الشبكة

احدى المشاكل التي تواجه المصمم لشبكة عصبية هي مدى عمومية شبكته ويمكن تلخيص هذه المشكلة كالتالي: بعد اكمال عملية تصميم وتدريب الشبكة العصبية وانخفاض معدل الخطأ الى أقل قيمة ممكنة وتمكن الشبكة من التعرف على كل العينات المستخدمة في التدريب فانه عند اختبار الشبكة على عينات جديدة للدخل نجد أن الشبكة تفشل في التعامل مع هذه العينات وتعطي نسبة اداء سيئة للغاية. بغرض التعامل مع هذه المشكلة وجعل الشبكة عامة لاي عينة دخل (تحسين عمومية الشبكة) يتم اتباع احد أسلوبين [4]:

(1) التنظيم (Regularization) :

هنا تدرب الشبكة حتى يصير اداؤها غير قابل للتغير وعندها فأن نسبة مجموع مربع الخطأ (sum squared error) للشبكة وكذلك مجموع مربع الأوزان (sum squared weights) تثبت عند اخر قيم وصلت اليها.

(2) الايقاف المبكر (Early stopping) :

هنا تقسم بيانات التدريب الى ثلاث مجموعات، المجموعة الاولى لتدريب الشبكة، والثانية لاختبار مصداقيتها في التعرف (validation) ، أما المجموعة الثالثة فتستخدم لاختبار مدى عمومية الشبكة. ثم تتم مراقبة الخطأ في المجموعة الثانية اثناء عملية التدريب فاذا فقدت الشبكة مصداقيتها في التعرف على العينات فان قيمة الخطأ هنا يزيد الى حد كبير وعندها يتم ايقاف عملية التدريب ويتم تثبيت الاوزان عند اخر قيمة وصلت اليها قبل أن تفقد الشبكة مصداقيتها. أما مجموعة الاختبار فلا تستعمل اثناء التدريب وانما فقط لاختبار اداء الشبكة عند النقطة التي تم تثبيت الأوزان عندها.

مشكلة البحث

في منظومة التعرف على الكلام نجد أن نظام اكتساب الأمر (هاتف، مايكروفون) يتأثر بالضجيج المحيط الذي قد يصل إلى درجة لا يمكن معها تمييز كلام المستخدم خاصة إذا كان الاكتساب يتم عن بعد كما في أنظمة التحكم الصوتي مع تحرير الأيدي. وهنالك مشكلة أخرى وهي أن نظام التحكم الصوتي يتأثر بالأشخاص اللذين لديهم عيوب في النطق وكذلك بالأشخاص الذين لا يتركون فاصلاً زمنياً بين كل منطوقة وأخرى. يمكن تلخيص مشكلات التعرف على الكلام كالتالي :

1 – الضجيج الخلفي (Back ground noise).

2 – عيوب النطق.

3 – الكلام السريع.

لذا فان تصميم شبكة عصبية بهيكلة قوية مع اختيار كافة البارمترات بعناية ومراعاة تحسين عمومية الشبكة قد يسهم بشكل واضح في التغلب على هذه المشكلات.

دراسات سابقة

قدمت الدراسة [7] تصميم لنظام تعرف على الكلمة المعزولة (isolated word recognition) باستخدام الشبكات العصبية الاصطناعية، وقد درب النظام على خمس وعشرين كلمة مختلفة تم نطق كل منها 30 مرة (15 مرة للتدريب ، و15 للاختبار). استخدم البحث تقنية مقاطعة الصفر (zero-crossing) للكشف عن نقاط البداية والنهاية لكل كلمة (End Points Detect)، وبغرض الحصول على استخلاص جيد للمعالم (Features Extraction) استخدم البحث أطوال مختلفة للاطر (Frames) ونسب مختلفة للتداخل بينها. وبالنتيجة فان أفضل طول لاطار كان 35ms وأفضل نسبة تداخل هي 0.6.

ثم استخدمت الدراسة شبكة عصبية ذات تغذية مرتدة من ثلاث طبقات (Back Propagation network) لتعمل كمتعرف على الكلام 25واستخدمت دالة التفعيل اللوغريثمية (Sigmoid logistic). واستخدم التدريب الموجة (supervised training) لتدريب الشبكة، وقد حققت النظام نسبة اداء % 94.4.

اشتملت الدراسة [8] تصميماً لشبكة ذات تغذية امامية متعددة الطبقات تستعمل خوارزمية الانتشار المرتد للخطأ في تنفيذ عملية التعرف على الارقام. حسث تم تقسيم الكلمات لاطر بطول 32 per 256 samples) ms frame) واستخدامت خوارزمية LPC (Linear Predicative Coding) بعد ذلك لاستخلاص معالم الكلمة حيث تم استخلاص 8 باراميترات من كل اطار وتم اختيار اربع اطر فقط لتمثل الكلمة.

الشبكة العصبية المستخدمة لها طبقة مخفية واحدة و10 مخارج تمثل الارقام من 0 – 9 وتستعمل الدالة اللوغريثمية لتفعيل خرج الطبقة، حققت الشبكة العصبية نسبة اداء وصلت الى 95% .

2.تصميم شبكة التعرف على الكلام

صممت منظومة التحكم الصوتي هنا لتتعرف على أمرين يشكل كل منهما كلمة واحدة وهما:

- افتح.
- اقفل.

وقد اشتمل تنفيذ هذا العمل على الخطوات التالية:

أولاً: تم تسجيل 55 عينة صوتية من كل أمر بصوت متحدث واحد فقط لتكوين قاعد بيانات النظام.

ثانياً: تم استدعاء كافة العينات الصوتية المكونة لقاعدة البيانات بواسطة برنامج المعالجة الأولية وقد استخدام برنامج MATLAB لكتابة هذا البرنامج وهو يحوي دوالاً صممت لاداء كافة العمليات الأساسية التي تشتمل عليها مرحلة المعالجة الأولية للكلام حيث يقوم أولاً باستقبال الاشارة الصوتية وتهيئتها وترشيحها من الضجيج المصاحب، بعد الترشيح تم تقسيم كل عينة الى أطر (Frames) في فترات صغيرة طول كل اطارms 20 وذلك لان إشارة الكلام تحوي العديد من الترددات والمعالم (Features) ولكي يجري استخلاص جيد لهذه المعالم فانه يجب أولاً تقسيم الإشارة لعدد من الأطر يتراوح طول الإطار الواحد ما بين 25 ms ~ 10 [1]، بعد ذلك تم ضرب كل اطار في window من النوع Hamming وقد استخدم النوع Hamming لانه الأنسب لتطبيقات التعرف على الكلام والغرض الاساسي من هذه العملية هو تنعيم شكل الموجة [1]، ويتم اختيار طول window بما يتناسب مع طول

الاطار، ولاطار بطول 20 ms فان طول window المناسب هو 30 ms [1]، وهنالك تداخل (Overlap) بين هذه والغرض من هذا التداخل هو تقليل مستوى الضجيج المصاحب. نسبة التداخل بين الاطر يمكن تحديدها باستخدام العلاقة التالية[1] :

$$\% \ overlap \ = \ \left(\frac{Tw \ - \ Tf}{Tw} \right) \ \times \ 100 \ \%$$

حيث Tw هي طول Window .

Tf تمثل طول الاطار.

في هذا العمل كان مقدار التداخل بين الاطر هو 66% من طول الاطار ، ثم استخرجت من كل اطار بارامترات LPC ، عادة يتم اختيار عدد 8~12 معامل من الاطار الواحد هنا تم اختيار 12 معامل واستخرجت منها معاملات Cepstral ، عدد معاملات cepstral يحدد كالتالي :

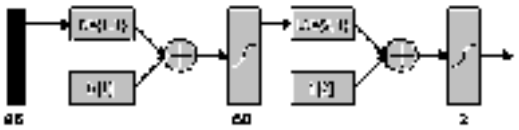
$$0.75 P \ \leq \ Nc \ \leq \ 1.25 P$$

حيث P هي عدد بارمترات LPC .

في هذا العمل تم اختيار 12 معامل Cepstral واستخرج منها متجة كتاب شفرة بطول 96 عنصر باستخدام خوارزمية Kmean [2]. هذا المتجة كان هو الدخل لشبكة التعرف على الكلام.

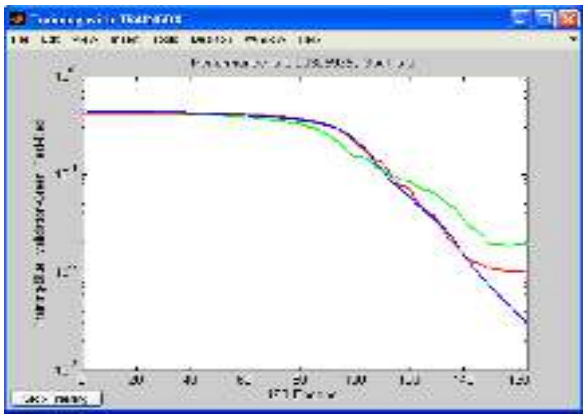
ثالثاً: تم تكوين شبكة عصبية لها 96 دخل وخرجين يمثل تفعيل أحدهما في كل مرة أمراً صوتياً معيناً. وقد تم اختيار شبكة عصبية ذات تغذية مرتدة وذلك لأن هذه البنية مناسبة جداً لتطبيقات التعرف على الكلام، أما بالنسبة لبنية الشبكة العصبية وعدد الطبقات والعقد المخفية فيها وكذلك خوارزمية التفعيل لكل طبقة فقد تم تصميمها بناء على الأسلوب المقدم في الفقرة (1) وقد كانت خطوات ونتائج التصميم كما يلي:

أولاً تم اختيار طبقة مخفية واحدة بين طبقتي الدخل والخرج بها 60 عقدة مخفية كما في الشكل (1).



شكل (1)

ولنفس بنية الشبكة العصبية تم تغيير خوارزمية التفعيل بغرض مقارنة الاداء كما يلي:



شكل (2): النتائج بعد استخدام شبكة عصبية بها طبقة مخفية واحدة

يوضح الشكل (2) النتائج بعد استخدام شبكة عصبية بها طبقة مخفية واحدة ولها بنية كالتالي:

عدد العقد = 60 ، الدالة Logsig

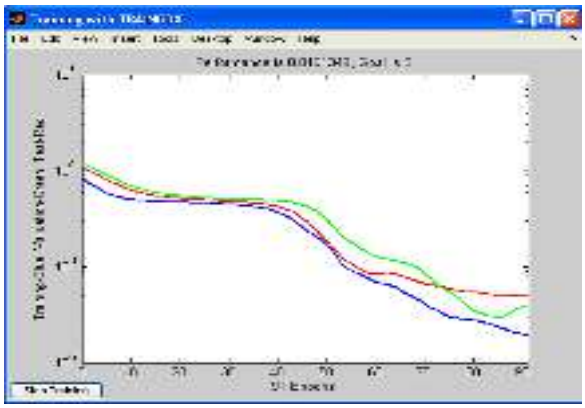
أما طبقة الاخراج فكانت كالتالي :

عدد العقد = 2 ، الدالة Logsig

يمثل الخط ذو اللون الأزرق االتناقص في الخطأ بالنسبة لعينات التدريب.

يمثل الخط ذو اللون الأخضر تناقص الخطأ في بينات كشف المصداقية مع عملية التدريب.

أما الخط الأحمر فهو خطأ الشبكة في التعرف على بيانات الاختبار.



شكل (3): النتائج بعد استخدام شبكة عصبية دالة تفعيلها Tansig

يوضح شكل (3) النتائج بعد استخدام شبكة عصبية بها طبقة مخفية واحدة كالتالي:

عدد العقد = 60 ، الدالة Tansig

أما طبقة الاخراج فكانت كالتالي :

عدد العقد = 2 ، الدالة Tansig

شكل (4): النتائج بعد استخدام شبكة عصبية دالة تفعيلها Purline

يوضح الشكل (4) النتائج بعد استخدام شبكة عصبية بها طبقة مخفية واحدة كالتالي:

عدد العقد = 60 ، الدالة Purline

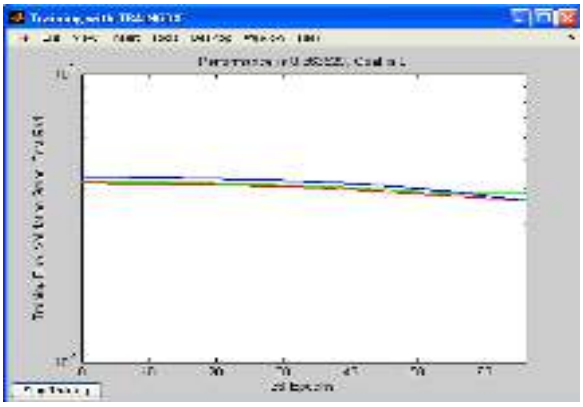
عدد العقد = 60 ، الدالة Purline

أما طبقة الاخراج فكانت كالتالي :

عدد العقد = 2 ، الدالة Purline

لاحظ تأثير تغيير دالة الانتقال (خوارزمية التفعيل) على اداء الشبكة حتى بدون تغيير عدد هيكلة الشبكة. من النتائج أعلاه من الواضح أن دالة Log sigmoid هي الأنسب لهذا التطبيق.

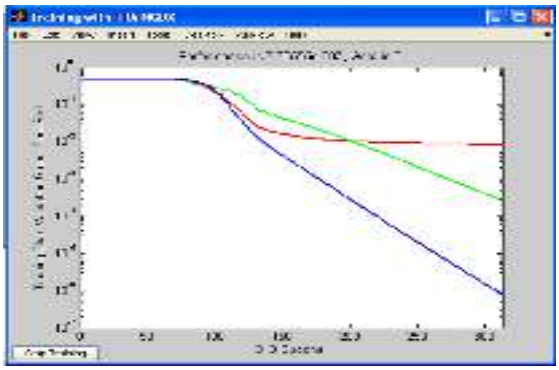
الان تأتي مرحلة تحديد عدد العقد المخفية، حيث تم تغيير عدد العقد من 60 الى 40 ثم الى 80 بغرض تحديد ما اذا كان التطبيق يستلزم زيادة العقد أم يستلزم انقاصها.



شكل (5): النتائج بعد تقليل عدد العقد الخفية الى 40 عقدة

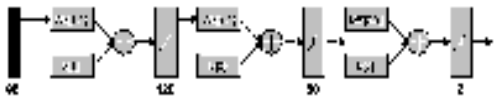
يمثل الشكل (5) النتائج بعد استخدام شبكة عصبية بها طبقة مخفية واحدة بها 40 عقدة.

من الواضح أن تقليل عدد العقد المخفية قلل كثيراً من مستوى الاداء اذن ننظر في اتجاه الزيادة ونزيد العقد الى 80 عقدة مخفية.



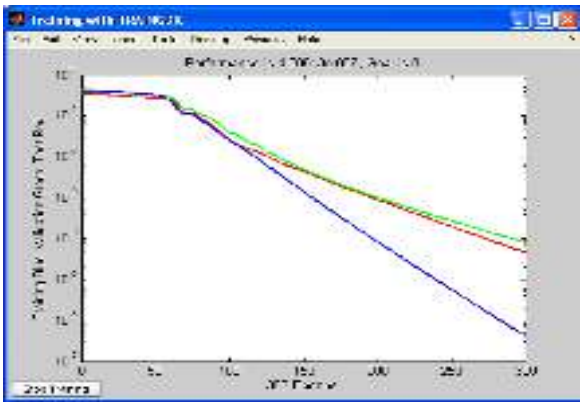
شكل (6): النتائج بعد زيادة عدد العقد الخفية الى 80 عقدة

الواضح أن الزيادة في عدد العقد ساهمت كثيراً في تحسين الاداء ولكن الخطأ في عينات الاختبار ما زال كبيراً. اخيراً تمت دراسة تأثير اضافة طبقة مخفية اخرى ليتم التحصل على الشبكة النهائية المستخدمة في التجربة والموضحة بالشكل (7) .



شكل (7) : الشبكة العصبية المستخدمة في التجربة الأولى

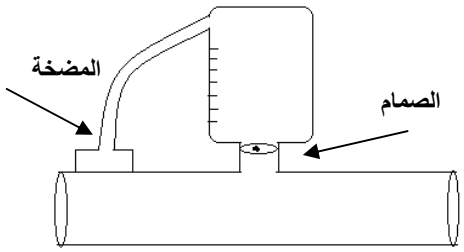
أما نسبة الخطأ في بينات التدريب وكشف المصداقية والاختبار فهو كما موضح بالشكل (8).



شكل (8): معدلات الخطأ في بينات التدريب وكشف المصداقية والاختبار

3 .تطبيق الشبكة العصبية في منظومة تحكم بالضغط :

تعاني انابيب توزيع المنتجات النفطية من ازدياد نسبة الضغط فيها وذلك بسبب الارتفاع في درجات الحرارة ولابد من وجود الية لمعالجة هذا حتى لا تنتج عواقب وخيمة هذه المنظومة تعمل على دفع قدر من المنتج الى خزان اضافي لتقلل من حجم الضغط على الانبوب ويتم استرجاع هذه الكمية مرة اخرى عند عودة الضغط الى الوضع الطبيعي الآمن، صممت هذه المنظومة لتعمل بشكل آلي والذي يجري عملية التحكم هنا هي شبكة عصبية صممت لتتخذ القرار المناسب في كل مرة فعندما يزداد الضغط تعطي امراً للمضخة بالعمل لدفع قدر من المنتج للخزان وعندما يستقر الضغط لاحقاً تعطي الشبكة الامر لصمام الكتروني موضوع بين الانبوب والخزان بالفتح لاستعادة المنتج من الخزان [3]. المنظومة آنفة الذكر موضحة بالشكل (9) أدناه:



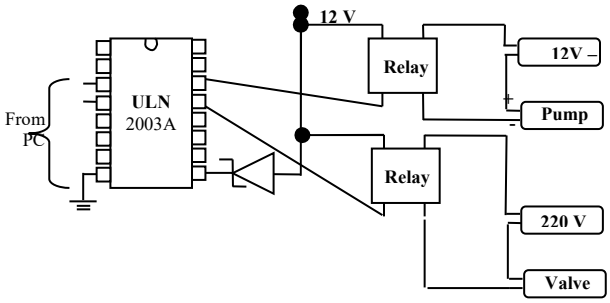
الشكل (9) : منظومة التحكم بالضغط

تم العمل على هذه المنظومة بحيث يمكنها أيضاً أن تستقبل أوامر صوتية وذلك في حالة الخطر الشديد عندما يقرر المهندس أن حالة الضغط الان شبة حرجة ومن الافضل للمنظومة أن تبدا العمل في هذه اللحظة على الفور عوضاً عن انتظار نقطة التشغيل التي دربت عليها الشبكة العصبية أو عند أي حالات مماثلة لذا فقد تم بناء شبكة عصبية لتتعرف على أمرين وهي التي يحتاج النظام لها وهي :

1. أفتح .
2. أقفل .

4. نمذجة النظام :

استخدم برنامج MATLAB في تحليل بيانات تدريب الشبكة العصبية واجراء المعالجة الأولية لها، بعد أكتمال تدريب الشبكة العصبية تم بناء برنامج التحكم الكامل الذي يستدعي الشبكة العصبية كدوال بداخله ، وقد استخدم MATLAB لكتابة هذا البرنامج . احتوى البرنامج ايضاً على دوال للموائمة (Interfacing) بين الحاسب والنظام أما دائرة الموائمة فهي كما في الشكل (10) .



شكل (10) : دائرة الموائمة

5. النتائج:

تم اختبار اداء المتعرف باستخدام عينات صوتية خارج نطاق العينات التي درب عليها، حيث تم تسجيل 20 عينة من كل أمر، وقد حقق النظام نسبة اداء 98.5714%.

6. الخاتمة والتوصيات:

تعد الشبكات العصبية تقنية قوية ومتطورة كما تمتاز بسهولة الاستخدام، واذا ما تم مراعاة تجويد تصميم الشبكة العصبية وفقاً للتطبيق المراد وذلك باختيار أفضل بنية وأفضل خوارزمية تفعيل وكذلك أمثل عدد للعقد الخفية فان ذلك يكسب الشبكة العصبية الاصطناعية وثوقية عالية وذلك مطلب هام في التطبيقات الحساسة مثل نظام التحكم بالضغط المقدم في هذه الدراسة. نسبة اداء المنظومة التي تم التحصل عليها هي 98.5714 % وهي نسبة مقبولة وتعطي وثوقية عالية للنظام.

يمكن تطوير المنظومة باضافة أوامر تتحكم بشكل منفصل في مكونات المنظومة مثل الصمام والمضخة. كما يمكن تحسين النتائج العملية باستخدام هجين بين تقنيتي الشبكات العصبية الاصطناعية ونموذج ماركوف الخفي (HMM-ANN).

المراجع :

[1] Joseph W. Picone, "Signal Modeling Techniques in Speech Recognition", IEEE, 1993.

[2] Manal Mohd. El. Obaid, Arabic Speech Recognition Using Neural Networks, 2005.

[3] Mohammed Kh. Hassen, Dalia M. Adam, Artificial Neural Networks and it's Application in Control Systems,2004.

[4] Mathwork.com, MATLAB Help Documents, Version. 7.0.0,2004.

[5] Yu Hen Hu, Jenq-Neng Hwang, Handbook of Neural Network Signal Processing, 2001.

[6] Eltaher M. Hussein, Dalia M. Adam, "Brain Tumor Detection Using Artificial Neural Networks", Journal of Science and Technology, Sudan University of Science and Technology, 2012.

[7]G. A bd ELAzeem , Ibrahim ELNahry and Amany Mahmoud Farg, Speech Recognition using Neural Networks, Mansoura Journal for Computer Science and Information Systems, Volume 1, 2005

[8] Chin Luh Tan and Adznan Jantan , DIGIT RECOGNITION USING NEURAL NETWORKS, Malaysian Journal of Computer Science, Vol. 17 No. 2, December 2004.