Data Classification using Quantum Neural Network

Dr. Ghassan H. Abd Al-Majeed  
Professor.  
Computer center- Baghdad University  
Baghdad, Iraq  
Email: ghmajeed@gmail.com

Dr. Zainab T. Alisa  
Assistant Professor.  
Electrical Engineering- Baghdad University  
Baghdad, Iraq  
Email: zainab_alisa@yahoo.com

Hassan Saadallah Naji  
Department of Electrical Engineering  
College of Engineering  
Baghdad University  
Email: hassanmaster79@yahoo.com

ABSTRACT

In this paper, integrated quantum neural network (QNN), which is a class of feedforward neural networks (FFNN’s), is performed through emerging quantum computing (QC) with artificial neural network(ANN) classifier. It is used in data classification technique, and here iris flower data is used as a classification signals. For this purpose independent component analysis (ICA) is used as a feature extraction technique after normalization of these signals, the architecture of (QNN’s) has inherently built in fuzzy, hidden units of these networks (QNN’s) to develop quantized representations of sample information provided by the training data set in various graded levels of certainty. Experimental results presented here show that (QNN’s) are capable of recognizing structures in data, a property that conventional (FFNN’s) with sigmoidal hidden units lack. In addition, (QNN) gave a kind of fast and realistic results compared with the (FFNN). Simulation results indicate that QNN is superior (with total accuracy of 97.778%) than ANN (with total accuracy of 93.334%).

Keywords: signal classification, artificial neural network, quantum computing, data analysis and fuzziness.
Data Classification using Quantum Neural Network

Ghassan H. Abd Al- Majeed
Zainab T. Alisa
Hassan Saadallah Naji

1. INTRODUCTION
Quantum neural network (QNN's) is a promising area in quantum computation and quantum information field. In 1997, Lov K. Grover proposed a method that can speed up a range of search applications over unsorted data using Quantum mechanics, Lov K. Grover, 1997.

Several models have been proposed in the literature but most of them need clear hardware requirements to implement such models, one of the most exciting emerging technologies is quantum computation, which attempts to overcome limitations of classical computers by employing phenomena unique to quantum-level events, such as nonlocal entanglement and superposition. It is therefore not surprising that many researchers have conjectured that quantum effects in the brain are crucial for explaining psychological phenomena, including consciousness, Abninder, 2006.

Jarernsri. L. Mitrpanont, and Ananta Srisuphab, presented the approach of the quantum complex-valued backpropagation neural network or QCBPN, the challenge of their research is the expected results from the development of the quantum neural network using complex-valued backpropagation learning algorithm to solve classification problems, Jarernsri, 2003.

Independent component analysis (ICA) is essentially a method for extracting useful information from data. It separates a set of signal mixtures into a corresponding set of statistically independent component signals or source signals. ICA belongs to a class of blind source separation (BSS) methods for separating data into underlying informational components, Isabelle, 2006.

The mixtures can be sounds, electrical signals, e.g., electroencephalographic (EEG) signals or images (e.g., faces, and Functional Magnetic Resonance Imaging (FMRI) data). The defining feature of the extracted signals is that each extracted signal is statistically independent of all the other extracted signals, James, 2004.

2. METHODOLOGY FOR INTEGRATED QNN SIGNAL CLASSIFIER SYSTEM

The overall block diagram that shows the structure of integrated QNN as signal classifier system is shown in Fig.1. Every single recorded input signal Iris signals database is depicted or formed by [1X50] discrete data matrix, and represents a vector pattern. Two different Iris data sets are formed for training and testing purposes. The discrete dataset has three different classes species, the structure of integrated QNN signal classifier system can be shown by the following principal steps:

2.1 Normalization
Normalization is a process to simplify data as feature extraction. It is usually affected by peak-to-peak magnitudes and offset of input data because of physiology conditions surrounding, psychological state, artifacts; therefore, normalization mainly required to decrease the effects of undesirable parameters and offset.

 الكرم المفتاحية: تصنيف الاشارات،الشبكة العصبية الصناعية،الإحصاء الكمي، تحليل البيانات، و المنطقة المضب.
2.2 Feature Extraction

Feature extraction or dimensionality reduction is the process of extracting useful information from the signal, features are characteristics of a signal that are able to distinguish between different classes species. Feature extraction requires reducing the size of the data by selecting appropriate features, selected features should be minimally redundant and the expected results should maximally depend on these features, and preserve all information from the signal that is needed for classification.

In ICA, each signal is described as a scalar variable, and a set of signals as a vector of variables, and the process of obtaining signal mixtures from signal sources using a set of mixing coefficients, Isabelle, 2006.

\[ x_1 = a_{11}s_1 + a_{12}s_2 \]  
(1)

\[ x_2 = a_{21}s_1 + a_{22}s_2 \]  
(2)

Above equations can be rewritten using matrix –by-vector form as:

\[
\begin{bmatrix}
X_1 \\
X_2
\end{bmatrix}
= \begin{bmatrix}
X_1 \\
X_2
\end{bmatrix}
\]

\[
S = \begin{bmatrix}
s_1 \\
s_2
\end{bmatrix}
\]

\[
A = \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\]

Then Eq. (1) can be written in matrix by vector form as follow:

\[ X = A \ S \]  
(6)

Where, \((a_{11}, a_{12}, a_{21}, a_{22})\), a set of mixing coefficients, \((s_1, s_2)\) are original signals(source signals), and \((x_1, x_2)\) set of “mixture” points which can be transformed back to the source signals \((s_1, s_2)\) using a set of unmixing coefficients, which reverse the effects of the original geometric transformation from source signals to signal mixtures, Joshua, 2000.

\[ s_1 = \mu x_1 + \sigma x_2 \]  
(7)

\[ s_2 = \gamma x_1 + \delta x_2 \]  
(8)

Above equations can be rewritten using matrix –by-vector form as:

\[
\begin{bmatrix}
s_1 \\
s_2
\end{bmatrix}
= \begin{bmatrix}
\mu & \sigma \\
\gamma & \delta
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\]

\[ W = \begin{bmatrix}
\mu & \sigma \\
\gamma & \delta
\end{bmatrix}
\]

Then Eq. (9) can be rewritten using matrix –by-vector form as:

\[
\begin{bmatrix}
s_1 \\
s_2
\end{bmatrix}
= \begin{bmatrix}
\mu & \sigma \\
\gamma & \delta
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\]

(9)

(10)
Data Classification using Quantum Neural Network

Ghassan H. Abd Al-Majeed
Zainab T. Alisa
Hassan Saadallah Naji

\( S = W X \) \hspace{1cm} (11)

Where \((\mu, \sigma, \gamma, \delta)\), a set of unmixing coefficients.

2.3 QNN and its Learning Algorithm

FFNNs must use the sample information as a mere reference for creating the internal representations, thus, it should not encode the sample information accurately into the internal representations. Such an exact or faithful encoding of the sample information results in the FFNN memorizing the “crispness” in the training data set. But an inherently fuzzy architecture should be capable of generalizing the sample information into various graded levels of certainty over the entire feature space. This may be possible if the architecture is capable of creating graded internal representations from the sample information. The QNN is as architecture capable of allowing the sample information to be encoded into certain levels grades of certainty/uncertainty only, Jarernsri, 2003.

One simple way of incorporating the ability to form consistent multilevel partitions in the hidden layer is to create hidden unit partitions with the property of “spreading-out” over regions of uncertainty in the feature space and collapsing-in over regions of certainty. If all the hidden unit transfer functions have the ability to form “graded” partitions instead of the crisp linear partitions, then these partitions can be “collapsed-in” or “spread-out” as required, using a suitable algorithm. Such an algorithm will not require that the fuzzy measures on the feature space be known, but will be a general procedure for learning the imprecision and uncertainty in the data set. This motivates the study of hidden units with multilevel transfer functions, Jarernsri, 2003 and Gopathy, 1996.

Suppose the multilevel hidden unit has \(n_s\) discrete states or levels. Then its transfer function can be written as a superposition of \(n_s\) sigmoidal functions, each shifted by \(\theta^f\). The output of this multilevel unit can be written as \(\frac{1}{n_s} \sum_{r=1}^{n_s} \text{sgm}(v^T x - \theta^r)\) where \(v\) is the connected weight matrix between input and the hidden units in hidden layers, and \(x\) is the input feature vector. The step widths of the multilevel transfer function, which may be called the quantum intervals, will be representative of discrete localized cells in the feature space consisting of feature vectors with approximately the same level of uncertainty as to their membership to the classes in the data set. These quantum intervals “jump-positions” \(\theta^f\) unlike the step widths, step heights need not be learned by independent parameters because several sigmoid can be shifted to the same location and added together to give steps of desired heights, to an approximation which reduces total number of parameters to be learned by almost one-third, Gopathy, 1996.

QNN consists of \(n_i\) inputs, one hidden layer of \(n_h\) nodes, each one represents a multilevel units and \(n_o\) outputs. Output units can be linear or sigmoid. Let \(x_k=[x_{1,k}, x_{2,k}, \ldots, x_{n,k}]^T, \ k=1,\ldots,m\), be the \(k^{th}\) feature vector of the data set \(X\). Then the input to the \(j^{th}\) hidden unit from the \(k^{th}\) feature vector is:

\[
h_{j,k} = \sum_{l=0}^{n_l} v_{jl} x_{1,k}
\]

\[
h_{j,k}^- = \frac{1}{n_s} \sum_{r=1}^{n_s} h_{j,k}^r = \frac{1}{n_s} \sum_{r=1}^{n_s} \text{sgm}(\beta_r(h_{j,k}^- - \theta^r))
\]

39
Where \( h_{j,k}^- \) : the response of the \( j^{th} \) multilevel hidden unit from \( k^{th} \) is feature, \( \text{sgm}(\tau) = \frac{1}{1 + e^{-\tau}} \) : is a sigmoid function (unipolar), \( \beta_h \) a slope factor for all multilevel hidden units in hidden layer, \( \theta_{j} \) define the jump positions in transfer function, and \( n_s \) is the number of levels or sigmoids in the multilevel hidden unit. Fig. 2 (a) plots the response \( h_{j,k}^- \) of \( j^{th} \) four level quantum neuron as a function of its input \( h_{j,k}^- \) with equal step heights and Fig. 2 (b) demonstrates multilevel transfer function with unequal step heights through simple shifting (Jarernsri, 2003).

\[
y_{l,k} = \sum_{j=0}^{n_{h}} w_{lj} h_{j,k}^-
\]

(14)

With \( h_{0,k}^- = 1 \), \( \forall k \) therefore, the response of the \( l^{th} \) output unit for the \( k^{th} \) input feature vector can be written as:

\[
y_{l,k}^\wedge = \text{sgm}(\beta_o (y_{l,k}^-))
\]

(15)

The major steps of QNN learning algorithm are summarized and presented as follows:

A) Update the synaptic weights:
Step 1: Select \( a_0, \beta_h, \beta_o \) (by trial and error) and randomly initialize the weights (\( W \) & \( V \)) and values of jump positions \( \theta_{j} \).
Step 2: Present \( k^{th} \) input pattern and specify the desired output.
Step 3: Calculate actual output \( y_{l,k}^\wedge \), using the present values of \( V_{ij} \) and \( W_{ij} \).
Step 4: Find the error terms (\( e_{l,k}^o \) and (\( e_{l,k}^h \)).
Calculate the output error.

\[
e_{l,k}^o = (y_{l,k}^\wedge - y_{l,k}) y_{l,k}^\wedge (1 - y_{l,k}^\wedge)
\]

(16)

Calculate the hidden layer error term.

\[
e_{j,k}^h = (\frac{1}{n_s} \sum_{r=1}^{n_s} h_{j,k}^r (1 - h_{j,k}^r)) \sum_{p=1}^{p=1} e_{p,k}^o w_{pj}
\]

(17)

\( p \) is over all nodes in the layer above node \( j \).
Step 5: Adjust the synaptic weights:

\[
w_{ij,k} = w_{ij,k-1} + \alpha e_{l,k}^o \frac{1}{n_s} \sum_{r=1}^{n_s} \text{sgm}(\beta_h (h_{j,k}^- - \theta_{j}^r))
\]

(18)

\[
v_{il,k} = v_{il,k-1} + \alpha h_{j,k}^o \frac{1}{n_s} \sum_{r=1}^{n_s} h_{j,k}^r (1 - h_{j,k}^r) x_{l,k}
\]

(19)

Where, \( k+1 \), \( k \), and \( k-1 \) index next, present, and previous respectively, and \( \alpha \) is a learning rate.
Step 6: Present another input pattern and go back to step 2
All of the training samples are presented cyclically.

B) Update the quantum intervals:
Step 1: In each training cycle, calculate the outputs $y_{ik}$, and $v_{jk}$ for each hidden node.
Then, take the average values for each class, $\langle h_{j,c_m}^{-}\rangle$, $\langle v_{j,c_m}^{s}\rangle$, for $m^{th}$ class during the training of QNN is as:

$$\langle h_{j,c_m}^{-}\rangle = \frac{1}{|c_m|} \sum_{x_k \in c_m} h_{j,k}^{-}$$  \hspace{1cm} (20)

$$\langle v_{j,c_m}^{s}\rangle = \frac{1}{|c_m|} \sum_{x_k \in c_m} v_{j,k}^{s}$$  \hspace{1cm} (21)

$|c_m|$ = The cardinality of $m^{th}$ class

Step 2: Calculate the quantum interval adjustment $\Delta \theta_q^r$ for each level:

$$\Delta \theta_q^r = \alpha_0 \sum_{n_s} \sum_{m=1}^{n_o} \sum_{x_k \in c_m} (\langle h_{j,c_m}^{-}\rangle - h_{j,k}^{-}) (\langle v_{j,c_m}^{s}\rangle - v_{j,k}^{s})$$  \hspace{1cm} (22)

Where $\alpha_0$ is the learning rate.

Step 3: Update jump-positions by:

$$\theta_j^r = \theta_j^r + \Delta \theta_q^r$$  \hspace{1cm} (23)

Step 4: Continue next cycle and go back to step 2.

Nomenclature:

- $n_o$ = number of output nodes.
- $n_i$ = number of input nodes.
- $n_s$ = number of quantum interval.
- $\theta_j^r$ = value of $r^{th}$ jump-position of hidden node $j$.
- $w_{ij}$ = the strength of connection between $j^{th}$ hidden node and $i^{th}$ output node.
- $v_{jl}$ = the strength of connection between $l^{th}$ input node and $j^{th}$ hidden node.
- $y_{ik}$ = actual output at output node $i$ for $k^{th}$ input pattern.
- $y_{ij}$ = desired output at output node $i$.
- $e_{ik}^{o}$ & $e_{ij}^{h}$ = error terms of output node $i$ and hidden node $j$ for $k^{th}$ input pattern.
- $h_{ik}^{o}$ = sgm ($h_{jk}^{r} - \theta_j^r$)
- $h_{jk}^{r} = \sum_{i=0}^{n_o} v_{jl} x_{ik}$, the internal state of hidden node $j$ for $k^{th}$ input pattern.
- $h_{j,k} = \frac{1}{n_s} \sum_{r=1}^{n_s} h_{j,kr}$, the output of $j^{th}$ hidden node.
\[ v_{j,k}^r = h_{j,k}^r (1 - h_{j,k}^r) \].

3. RESULTS AND DISCUSSION
3.1 Results of using ANN Classifier

Three layer ANN was employed as classifier, for Iris data set signal, network structure is 4-16-3 namely input vector \((n_i)\) is equal to \((4)\) representing number of input variables. The 4 rows (sepal length, sepal width, petal length, and petal width) contained a single hidden layer, number of hidden units was chosen by trial and error; which revealed that number of neurons at hidden layer \((n_h)\) is at least greater than twice of input nodes \((n_i)\) correspond to features of each species, as an assumption \((n_h = 4*n_i)\) for good performance \((n_h = 16)\) traditional neuron with sigmoid activation function (unipolar), and output layer contains number of neuron equivalent to classes (species), \((n_o = 3)\), structure of the ANN classifier is shown in Fig.3 with number of nodes at each layer.

Three layers ANN was employed as classifier of Iris data signal, randomly select input feature vectors to achieve the uncertainty principle. Randomly selected 70% samples as training samples and 30% for testing samples. Learning rate \((\eta)\) is chosen by trial and error for weight adjusting \((W_{ij}, V_{jl})\) is set to \((0.01)\) MATLAB programming test, the number of iteration (epochs) is set to 1500.

The performance of ANN classifier (for both training and testing phases) is shown in Fig.4 and for training data set; ANN gave a Mean Square Error (MSE) of \((0.0054)\) and accuracy of \((100.00\%)\).

For testing phase, ANN classifier showed a mean square error MSE of \((0.1174)\) with accuracy of \((93.334\%)\). Fig.4 shows the relationship between MSE and number of iteration for ANN classifier. As can be seen that MSE is decreasing with increasing number of iterations, which revealed that the network is converged with iteration number of \((1500)\) and MSE of \((0.0054)\).

3.2 Results of using QNN Classifier

Three layers QNN was employed as classifier of Iris data signal, the performance of the QNN was tested to perform classification on Iris data signal. Randomly selected 70% of the samples are used for training (training input feature vectors) and 30% for testing (testing input feature vectors), the structure is 4-8-3 for the neural network, input vector \((n_i)\) is equal to \((4)\).

The number of multi-level hidden units (which are used in the hidden layer of QNN rather than traditional neurons as in ANN) with the number of multi-level neurons \((n_h)\) is no more than twice of the input nodes \((n_i)\), in such structure of \((2*n_i)\), i.e. \((n_h = 8)\) multi-level hidden units with sigmoidal activation function (unipolar), and the output layer contains neurons equivalent to number of species \((n_o = 3)\). The major difference between QNN and ANN is that the QNN uses quantum neuron (multi-level neuron (graded) with sigmoid activation function), structure of QNN classifier is shown in Fig.5. It contains a single hidden layer with \((n_h = 8)\) units, three output units (no. of species) and 4 input nodes, also identifies by existence of jump-positions (thetas’ values) of multi-level hidden units of the QNN hidden units were chosen by trial and error.

The QNN is composed of multi-level hidden units with \((n_s = 3)\) (chosen by trial and error) levels for each hidden unit, the learning rate (learning ratio) \((\eta)\) for weight adjusting \((W_{ij}, V_{jl})\) is set to \((0.07)\) training by MATLAB test, and the learning rate for quantum interval adjusting \((\eta_theta)\) is set to \((0.001)\), and slop factor for unit at hidden layer \(\beta_h = 2\), but for output layer slop factor \(\beta_o = 1.5\), the number of iteration (epochs) is set to \((300)\). Synaptic weights \((W_{ij}, V_{jl})\) are adjusted by
minimizing quadratic error function with respect to particular weights, in training itself; jump-positions of multi-level hidden units adjusted also.

Results are done with (70%) of input feature vectors for training phase and (30%) of input feature for testing phase, performance of QNN classifier shown in Fig.6 and summarized with (1)MSE of (0.1258) and accuracy of (97.143%) for training phase (2) MSE (0.3780) with accuracy of (97.778%) for testing phase. Fig.6 displayed the relationship between MSE and number of iteration for QNN classifier, in which reveals that MSE is decreasing with increasing number of iterations, but when the iteration number exceeded 150 the MSE decreases very slightly which means the network is converged with iteration number of (150) and MSE of (0.1258). As in Fig.4, Fig.6 shows an inverse relationship between MSE and number of iterations.

Main difference between two figures is that QNN converges with less iteration, Fig.6 shows convergence occurs at (150) in QNN classifier compared with (1500) for ANN.

3.3 QNN vs. ANN Classifiers

To discuss the results of ANN and QNN classification for Iris data signals (QNN vs. ANN classifiers), table 1 shows a comparison between them. Two issues can be concluded from table1, first one is that QNN classifier gave better accuracy for testing phase compared with ANN classifier (97.778% compared with 93.334%) and second issue is that QNN converged with less number of iterations (150 epochs for QNN compared with 1500 epochs for ANN), and this indicates that the time required for QNN convergence is about (90%) less than that of ANN.

The reason is that ANN is unable to correctly assign class membership to data samples belonging to regions of the feature space where there is overlapping among the classes. The reason for this is that FFNNs use sharp decision boundaries (due to crisp membership function) to partition the feature space. As a result, the outputs of trained ANNs cannot generally be interpreted as membership values. Also it can be found that QNN is more reliable than the ANN because QNN generates a more structured representation of the input data at the hidden layer than that of the ANN as QNN use multilevel hidden units, this is not surprising, given the fact that the jump-positions of the multi-level hidden units of the QNN are updated by minimizing some measure based on the class-conditional variances at the outputs of the hidden units.

Another advantage is that QNN systems are using quantum neuron instead of traditional neuron which is often able to learn faster and require less number of neurons in the hidden layer which could lead to a smaller number of weights and reduction of the number of neurons in the hidden layer which could lead to smaller number of weights, or it can be said generally that this means reducing the total number of parameters (input weights, output weights, jump position) to be learned by almost half, as an assumption to the total number for hidden units in hidden layer is with(4*n1) empirically for best accuracy in ANN classifier, while assumption to the total number for multilevel hidden units in hidden layer is with(2*n1) empirically for best accuracy in QNN classifier. This means that the total number of parameters (input weights, output weights, jump position) to be learned is reduced by almost half which represents another advantage for QNN.

Here the number of multilevel hidden units in the hidden layer was set to (8) by QNN but (16) by ANN, the input weights is (V(8x4)) and output weight (W(3x8)) and jump position with (8x3)matrix by QNN while the input weight is (V(16x4)) and output weight is (W(3x16)) for ANN.
4. CONCLUSIONS

In this paper, a data classification system based on multi-level transfer function integrated Quantum Neural Networks (QNN) is proposed. The classification system methodology consists of ion signals which here iris data. First, the classification signals should be normalized then feature extraction is applied using independent component analysis technique then classification task is to be achieved firstly using artificial neural network classifier and secondly using integrated Quantum Neural Networks (QNN).

The architecture of (QNN’s) has inherently built in fuzzy, hidden units of these networks (QNN’s) to develop quantized representations of sample information provided by the training data set in various graded levels of certainty. Experimental results presented here show that (QNN’s) are capable of recognizing structures in data, a property that conventional (FFNN’s) with sigmoidal hidden units lack. In addition, (QNN) gave a kind of fast and realistic results compared with the (FFNN). Simulation results indicate that QNN is superior (with total accuracy of 97.778%) than ANN (with total accuracy of 93.334%).

REFERENCES


Figure 1. Block diagram of the methodology for signal classification system.
Figure 2. Multilevel transfers function with (a): equal step heights (b): unequal step heights.
Figure 3. Structure of the Artificial Neural Network (ANN) classifier.
Figure 4. Classification of signals by ANN classifier.
Figure 5. Structure of Quantum Neural Network (QNN) classifier.
**Figure 6.** Classification of signals by QNN classifier.

**Table 1.** Classification and performance using ANN and QNN classifiers.

<table>
<thead>
<tr>
<th>Network classifier type</th>
<th>MSE for training set</th>
<th>Accuracy for training set (%)</th>
<th>No. of epochs for convergence</th>
<th>MSE for testing set</th>
<th>Accuracy for testing set (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN classifier</td>
<td>0.0054</td>
<td>100.00</td>
<td>1500</td>
<td>0.1174</td>
<td>93.334</td>
</tr>
<tr>
<td>QNN classifier</td>
<td>0.1258</td>
<td>97.143</td>
<td>150</td>
<td>0.3780</td>
<td>97.778</td>
</tr>
</tbody>
</table>