Chapter from the book *Artificial Neural Networks - Application*

Downloaded from: http://www.intechopen.com/books/artificial-neural-networks-application

Interested in publishing with IntechOpen?
Contact us at book.department@intechopen.com
Application of Artificial Neural Networks in Optical Properties of Nanosemiconductors

M. Farzalipour Tabriz¹, P. Salehpour² and A. Esmaielzadeh Kandjani¹

University of Tabriz
¹Faculty of Mechanical Engineering, ²Faculty of Electrical Engineering and Computer, Iran

1. Introduction

1.1 Nanoscience

Nanoscience is an interdisciplinary field which deals with the unconventional phenomena observed in materials with at least one characteristic dimension in the range of 1-100 nm. Nanoscale materials exhibit some novel and/or improved properties over either atoms/molecules or bulk state which is resulted from limited size of their constituent components. There are two main issues discussed in nanomaterials. One of them is high surface to volume ratio for nanomaterials over bulk. In nanomaterials, higher percentage of atoms is located on the surface which leads to high specific surface area, as shown in figure 1.

![Fig. 1. Variations of the percentage of surface atoms by increasing the total number of atoms in close packed icosahedron clusters (adapted from geometrical model data in (Allpress & Sanders, 1970))](image)

Atoms at the surface have fewer direct neighbors than atoms in the bulk. Therefore, nanomaterials have a large fraction of their atoms at surface with a low average coordination number (which is the number of nearest neighbors), high surface energy and
diffusion rates (Jiang et al., 2004). Surface properties are of particular interest in this subject because of their importance in chemistry (Burda et al., 2005) and influence on electronic and optical properties (Puzder et al., 2002). Solid–gas or solid–liquid chemical reactions can be mostly confined to the surface and/or subsurface regions of the solid. The interior atoms in particles are more highly coordinated, form more bonds and therefore are more stable than those at the surface. For this reason, the surface (especially edge and corner) atoms normally exhibit the highest affinity to form bonds with other molecules. This fact is of utmost importance for chemical activity (Burda et al., 2005). Thus, particle size distribution and shape play important role in determining properties of nanomaterials. Another issue is quantum confinement effect which is change of electronic and optical properties of materials when their structural size is sufficiently small - typically 10 nanometers or less (Stucky & MacDougall, 1990).

1.2 Nanosemiconductors
Nanocrystalline semiconductors and specially their electrical and optical properties have been studied extensively in recent years (Pal, 1999; Schmitt-Rink et al., 1987). These materials behave differently from bulk semiconductors. With decreasing particle size the band structure of the semiconductor changes; the band gap increases and the edges of bands split into discrete energy levels. Specifically, the phenomenon results from electrons and holes being squeezed into a dimension that approaches a critical quantum measurement, called the exciton Bohr radius. But researchers have also found band gap increase in some nanomaterials having sizes far beyond the quantum confinement regime (Chen et al., 2006). The band gaps have a major influence on the properties of the semiconductors including optical absorption, electrical conductivity and index of refraction. According to the band gap type, semiconductors are normally classified as direct and indirect.

![Diagram of direct and indirect band gaps](image-url)

Fig. 2. a) Direct and b) Indirect transition of electrons in semiconductors.

When the carriers transfer between the conduction and valence band, in direct band gap the momentum (or k) is conserved but it’s changed for indirect band gap semiconductors. The direct band gap semiconductors are found to be advantageous over indirect band gap semiconductors, as they do not require phonons to satisfy crystal momentum conservation (Sze, 1981).
In semiconductors, the main characteristic of light absorption is slightly different from metals and insulators. This is due to their band structure where the band gap between lowest point of conduction and upper point of valance bands is low enough to allow the transferring of electrons from valance band to conduction band by exciting with energy sources like light, electrical field, etc. Thus, the energy difference between conduction and valance bands (i.e. band-gap) could be detected by considering the energy absorbance of semiconductors (Sze & Ng, 2007).

When a light beam propagates into a media rather than vacuum, based on its photon energy, a portion of light is absorbed, another part is transferred and a small portion could be reflected. The physical meaning of absorbance could be defined as the logarithmic proportion of intensity of propagating light with specified wavelength passed through a sample to the intensity of the light before entering it. The absorbance which is detected via UV-Vis spectroscopy always shows this absorbance which is defined as optical density (Zhang, 2009).

On the other hand, the light absorbance can be expressed by light absorption coefficient \( \alpha(h\nu) \), which is the relative decrease rate of propagating light intensity \( I(h\nu) \) passing through a matter along its propagating path \( x \) (Gaponenko, 1998):

\[
\alpha(h\nu) = \frac{1}{h\nu} \frac{dI(h\nu)}{dx}
\]  

(1)

Absorption coefficients of colloidal suspension \( (\alpha, \text{ cm}^{-1}) \) have been calculated using the following equation:

\[
\alpha = 2303 \left( \frac{D\rho}{C\ell} \right)
\]

(2)

where, \( D \) is the optical density of a solution which is measurable by UV-Vis absorbance spectroscopy, \( \rho \) is the density of dispersed particles, \( C \) is its concentration and \( \ell \) is the optical path (cm). The main interest in calculating \( \alpha(h\nu) \) is due to its relation with semiconductor’s band gap energy. This relation can be expressed as follow:

\[
\alpha(h\nu) = A' (h\nu - E_g)^n
\]

(3)

where \( A' \) is a constant determined by the index of refraction, and electron and hole effective masses, or

\[
\alpha(h\nu) = A \frac{(h\nu - E_g)^n}{h\nu}
\]

(4)

where \( E_g \) is semiconductor’s band gap \( A \) is a coefficient of the given electronic transition probability and \( n \) equals to 0.5 (for direct band gap semiconductors) and 2 (for indirect band gap semiconductors) in allowed direct and indirect transitions (Zhang, 2009).

When a photon with sufficient energy is absorbed with a semiconductor, an electron can be excited from valance band and move to conduction band and subsequently, an electron-hole pair is generated in the semiconductor. In this sense an optical excitation is a two-particle transition, an electron and a hole. Existing two different type particles (i.e. an electron and a
hole) makes a Columbus interaction between the electron and hole and thus it is act as a hydrogen atom. This electron-hole is namely known as exciton. Excitons are potentially mobile and, due to producing a positive hole by exciting a negative electron, are neutral charged (Klingshirn, 2005).

Excitons are divided into two general categories, Mott-Wannier excitons and Frenkel excitons. Mott-Wannier excitons have weak electron-hole interactions caused by a small Coulomb attraction due to relatively far distance between constituents. Corresponding binding energies are on the order of 10meV. By contrast, in Frenkel excitons the carriers are close and have strong Coulomb interactions. Corresponding binding energies are on the order of 100meV. Frenkel excitons are commonly seen in organic semiconductors while in nonorganic semiconductors Mott-Wannier excitons are the main detectable excitons (Klingshirn, 2005).

Excitons can be detected in the absorption spectrum of semiconductors as extrema points in absorption or transmission spectra. They generally appear just below the band edge of the semiconductor. This is because the energy of the exciton is lower than the band edge transition by its binding energy.

Although most oxides are good insulators but some of them such as ZnO and TiO$_2$ are well-known semiconductors. Oxide semiconductors are very interesting materials because they combine easily adjustable electronic properties with relatively high working temperature. Some examples of the oxide semiconductors (with their corresponding band gaps energy) are Cu$_2$O (2.1 eV), Bi$_2$O$_3$ (2.8 eV), TiO$_2$ (3.2 eV), ZnO (3.4 eV) and SnO$_2$ (3.7 eV), BaTiO$_3$ (3 eV), SrTiO$_3$ (3.3 eV) and LiNbO$_3$ (4 eV). These materials are employed in a variety of electronic applications, such as positive temperature coefficient thermistors (Goodman, 1963), varistors (i.e., resistors with nonlinear, but symmetric, current–voltage characteristics which are used for the protection of electronic devices and circuits) (Levinson & Philipp, 1975), capacitors of high dielectric constant (Robertson, 2004) and gas sensors (Eranna et al., 2004).

### 1.3 Zinc oxide

Among the studied metal oxide nanomaterials, zinc oxide is a notable case. Zinc oxide is a II-VI wide band-gap semiconductor, with a direct band gap of 3.37 eV (at room temperature in bulk state) and large exciting band energy (60 meV) (Madelung et al., 1999).

Thermodynamically stable crystalline structure of ZnO under ambient conditions is hexagonal wurtzite (space group P6$_3$mc) with a = 0.32501 nm and c = 0.52071 nm (Kisi & Elcombe, 1989). This structure can also be described as alternating stacking of O and Zn ionic planes along the c axis. Absence of inversion symmetry (center of symmetry) in this crystalline structure is the origin of its piezoelectric and pyroelectric properties (Hübner, 1973).

Wurtzite crystals are dominated by four low Miller index surfaces: the nonpolar (10$ar{1}$0) and (11$ar{2}$0) surfaces and the positively charged polar zinc terminated (0001)–Zn and the negatively charged oxygen terminated (000$ar{1}$)–O surfaces (Diebold et al., 2004). These polar surfaces results in a normal dipole moment and spontaneous polarization along the c axis as well as a divergence in surface energy which influences the adsorption of existing ions in reaction media which can affect the morphology (Wang, 2005) and properties (Jang et al., 2006) of resulted materials.

X-ray diffraction is one of the most important characterization tools in materials science. X-ray diffraction can be used for characterization of crystalline materials and the
Fig. 3. Stick-and-ball model of a) alternating stacking of O and Zn ionic planes along the c axis, and b) unit cell of wurtzite structure from the results of (Xu & Ching, 1993) determination of their structure. The diffraction of X-rays by a crystalline solid results in a pattern of sharp Bragg diffractions characteristic of the different d-spacings of a solid. A typical X-Ray diffraction pattern of wurtzite zinc oxide is shown in figure 4.

Fig. 4. X-Ray diffraction pattern of wurtzite zinc oxide (JCPDS No.036-1451)
Broadening of these reflections beyond that arising due to instrumental factors is generally attributed to crystallite size effects. The average crystallite sizes of particles can be estimated from the full width at half maximum (FWHM) of the highest X-ray diffraction peaks using the Debye–Scherrer equation (Cullity, 1978):

\[ D = \frac{k\lambda}{\beta \cos \theta} \]

where \( D \) is the mean crystallite size; \( k \) is a grain shape dependent constant (0.89 for spherical particles); \( \lambda \) is the wavelength of the incident beam; \( \theta \) is the Bragg diffraction peak angle; and \( \beta \) is the full width at half maximum.

Nanocrystalline ZnO is widely used in many applications such as blue light emitting diodes (Tsukazaki et al., 2005), photo detectors (Jun et al., 2009), gas sensors (Xu et al., 2000), photocatalysts (Hariharan, 2006), and field-effect transistors (Arnold et al., 2003). Various
chemical methods have been successfully used for synthesizing nanostructured ZnO material, such as hydrothermal (Xu et al., 2004), sol-gel processing (Ristic et al. 2005; Li et al., 2005), chemical bath deposition (Wu et al., 2006) and sonochemical synthesis (Kandjani et al., 2008).

1.4 Design of experiments and chemical synthesis
Increase in the demands for chemical goods with specific desired properties has made controlling of synthesis methods as a challenging scope in production of these goods. As the variables for a synthesis could vary from less than ten to the thousands based on the synthesis route and also the fact that most of synthesis parameters influence each other’s effects, controlling and obtaining the situations in which the desired properties could be achieved from synthesis has become one of the most important complexity for many researches (Box & Draper, 2007). Thus, for obtaining optimal properties of the final product a lot of experiments should be conducted in classical method. Although, the simultaneous effects of variables on final properties can’t be clearly determined due to exponentially increase in experiments needed for covering all features of the synthesis by increase in the numbers of the variables (Lazic, 2004). Thus, designing a method for decreasing costs and also time schedule for understanding a process has become almost crucial for high-tech systems where costs and speed in achieving the answers has a critical role in getting results and survival in technological competition.

To design an experiment means to choose the optimal experiment design to be used simultaneously for varying all the effective parameters. By designing an experiment one gets more precise data and more complete information on a studied phenomenon with a minimal number of experiments and the lowest possible cost. Two approaches are available for design of experiments; first, classical experimental design (one factor at a time-OFAT) and second, statistically designed experiments (DOE) (Ferreira et al., 2007). The latter which allow the simultaneous study of several control variables, are faster to implement and more cost-effective than traditional uni-variable approach.

There are two different conditions occurs on most of experiments, first-order and second-order models. When a linear function can be used to describe a phenomenon, first-order model is applicable. For data which do not obey simple linear functions or when an optimization is necessary, first-order models are not a suitable so the second-order models are usually used such as Box–Behnken design (BBD), central composite (CCD), and Doehlert (DD) designs. The efficiency of these designs decreases by increase of the variables. The efficiency decrease rate in Doehlert design is much slower than other designs by increasing variables. In the other words, among the mentioned second-order models Doehlert is the most efficient design (Ferreira et al., 2004). The main characteristic of this model could be summarized as (Bezerra et al., 2008):

- Total number of needed experiments are $N= k^2 + k + 1$, where k is the variables number;
- Each variable is studied at a different number of levels;
- The intervals between its levels make a uniform distribution; displacement of the experimental matrix to another experimental region can be achieved using previous adjacent points.

For three variables, DD is represented by a geometrical cuboctahedron, and, depending on how this shape is projected on the plane, it can generate different experimental matrices. Table 1 shows the experimental matrix of DD for a two variables experiment and two different matrices for DD with three-variable. The explanation of how these matrixes obtained is beyond the scope of this chapter.
Two Variables | Three Variables (Type 1) | Three Variables (Type 2)
---|---|---
| X₁ | X₂ | X₁ | X₂ | X₁ | X₂ | X₁ | X₂ | X₁ | X₂ |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | -1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 0.5 | 0.866 | 1 | 0 | 0 | 0.5 | 0.866 | 0 | 0 | 0 |
| -0.5 | -0.866 | -1 | 0 | 0 | -1 | 0 | 0 | 0 | 0 |
| 0.5 | -0.866 | -0.5 | -0.5 | 0.707 | -0.5 | 0.866 | 0 | 0 | -0.817 |
| -0.5 | 0.866 | 0.5 | -0.5 | 0.707 | -0.5 | -0.289 | -0.817 |
| 0.5 | 0.5 | 0.707 | 0.5 | -0.866 | 0 | 0 | 0 | 0 |
| -0.5 | 0.5 | 0.707 | 0.5 | -0.866 | -0.817 |
| -0.5 | -0.5 | -0.707 | 0 | 0.577 | -0.817 |
| 0.5 | 0.5 | -0.707 | 0.5 | 0.289 | 0.817 |
| -0.5 | 0.5 | -0.707 | 0 | -0.577 | 0.817 |

Table 1. Doehlert matrices for two and three variables

1.5 ANN application in nanotechnology
Statistical experimental designs have been widely used in nanoscience and technology to determine the combined effects of variables for the goal of optimizing desired properties but they need a mathematical model to estimate the results in the domain of interest. Unfortunately the underlying relations between the processing parameters and the properties of nanomaterials have not yet been fully understood so usually no analytical model is available as the relation of parameters under investigation and using empirical models often gives inaccurate results. For this purpose using artificial neural network is usually considered a more beneficial approach for modeling these poorly understood datasets of experimental results. Artificial neural networks have been reported to be successfully used for modeling of growth rate (Yo et al., 2009), grain size (Rashidi et al., 2009), particle size (Khanmohammadi et al., 2010), photocatalytic properties (Kandjani et al., 2010), magnetic properties (Mohorianu et al., 2009), oxidation kinetic (Straszko et al., 2008) and emulsion stability (Amani et al., 2010) of nanomaterials and also the effects of parameters in ball milling processing (Ma et al., 2009), sol-gel synthesis (Fan & Liu, 2009), spray reaction synthesis (Zhang et al., 2007) on resulted materials.

1.6 Hydrothermal synthesis
The term “hydrothermal” first was used by British geologist Sir Roderick Murchison in 1840s for describing the action of water at high pressures and temperature on earth crust which leads to natural formation of rocks and minerals. It’s now referred to heterogeneous reactions in aqueous solution or mineralization in high temperature and pressure. Hydrothermal process permits dissolution and recrystallization of materials which are not soluble under normal conditions. Thus, hydrothermal can be used in various processes i.e. crystallization (Matthews, 1976), crystal growth (Laudise & Nielsen, 1961), synthesis (Somiya & Roy, 2000), decomposition (Jomaa et al., 2003), extraction (Goguel, 1985), etc. Lots of researches in this field have been dedicated to “Hydrothermal synthesis” which usually involves aqueous chemical reactions at temperatures higher than 100ºC in a closed system.
This synthesis method in comparison with other conventional chemical methods has higher efficiency and controllability and its products have high purity and good crystalline quality with narrow particle size distributions.

Hydrothermal method has been successfully used for synthesizing a wide variety of nanomaterials with different morphologies which an extensive review can be found in (Byrappa & Adschiri, 2007).

There are a number of parameters which can influence the products of hydrothermal synthesis. Some of the most important variables include temperature, duration and concentration of reactants. The temperature determines the solubility of materials in water and the pressure in closed autoclaves and also influences the diffusion/reaction rate and Gibbs free energy changes for various reactions. Duration of synthesis provides needed time to reach the thermodynamically stable state of phase, size and morphology (in elevated temperature and pressure). Initial concentration of reactants determines the reaction rate and influences the nucleation and growth rate, reaction mechanism and final obtained phase.

One of the models, which can describe formation and growth mechanism of crystalline ZnO from solution, is growth unit model (Li et al., 1999). Zhong firstly presented the growth unit model in early 1990s (Zhong et al., 1991; Zhong et al., 1994). In this model, growth units are the polyhedral complexes with \((OH^-)\) ligands, in which the cations have the same coordination number as in the oxide crystal lattice. According to this model, the growth units of ZnO crystals are \(Zn(OH)_{2-}\) complexes which produce zinc oxide by sharing elements as following:

\[
Zn(OH)_4^{2-} + Zn(OH)_4^{2-} \rightarrow Zn_2O(OH)_6^{4-} + H_2O
\]  

(6)

These reactions yield ZnO particles with \(OH^-\) ligands on their surfaces. It has been reported that, Zn(OH)\(_2\) is predominantly formed at pH 6-9, while wurtzite ZnO is mainly obtained at pH 9-13 (Yamabi & Imai, 2002).

Drying of samples cause to Zn(OH)\(_2\) crystals to decompose into ZnO by forced hydrolysis (Matijevic, 1985):

\[
Zn \,(OH)_2 \rightarrow ZnO + H_2O
\]  

(7)

2. Materials preparation

Sodium hydroxide, NaOH and zinc acetate, Zn(O\(_2\)CCH\(_3\))\(_2\).2(H\(_2\)O) were purchased from Merck and were both used without further purification.

Doehlert experimental design was used for investigating the effect of synthesis temperature, synthesis time and initial concentrations of precursors on the properties of synthesized nanopowders. The ratio of the Zn(Ac)\(_2\) to NaOH was kept equal to 1/2. The complete experimental design and variables are listed in table 2.

First; aqueous solution of Zn(Ac)\(_2\) was added dropwisely to the same amount of NaOH aqueous solution with defined concentration. The obtained solution was poured into 35ml PTFE lined autoclaves (up to 80%Vol). Then the autoclaves were kept at defined temperature. After a defined period of time, autoclaves were cooled to room temperature naturally and the resulted precipitates were filtered and washed with distilled water and ethanol for several times. Finally obtained powders were dried at 50°C for 24 hours.
Table 2. Multivariate experimental design for present study

<table>
<thead>
<tr>
<th>Sample</th>
<th>Temp. [°C]</th>
<th>Time [hrs]</th>
<th>Zn(Ac)$_2$ [mol]</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>150</td>
<td>12:00</td>
<td>0.75</td>
</tr>
<tr>
<td>S2</td>
<td>190</td>
<td>12:00</td>
<td>0.75</td>
</tr>
<tr>
<td>S3</td>
<td>170</td>
<td>17:12</td>
<td>0.75</td>
</tr>
<tr>
<td>S4</td>
<td>170</td>
<td>13:44</td>
<td>0.5</td>
</tr>
<tr>
<td>S5</td>
<td>110</td>
<td>12:00</td>
<td>0.75</td>
</tr>
<tr>
<td>S6</td>
<td>130</td>
<td>6:48</td>
<td>0.75</td>
</tr>
<tr>
<td>S7</td>
<td>130</td>
<td>10:18</td>
<td>0.5</td>
</tr>
<tr>
<td>S8</td>
<td>170</td>
<td>6:48</td>
<td>0.75</td>
</tr>
<tr>
<td>S9</td>
<td>170</td>
<td>10:18</td>
<td>0.5</td>
</tr>
<tr>
<td>S10</td>
<td>130</td>
<td>17:12</td>
<td>0.75</td>
</tr>
<tr>
<td>S11</td>
<td>150</td>
<td>15:30</td>
<td>0.5</td>
</tr>
<tr>
<td>S12</td>
<td>130</td>
<td>13:44</td>
<td>1</td>
</tr>
<tr>
<td>S13</td>
<td>150</td>
<td>8:30</td>
<td>1</td>
</tr>
</tbody>
</table>

3. ANN modeling

In these studies, special software was developed to be able to compute accurate neural network results. This program was written in C++ and tried to be as fast as possible to create more results in shorter time. It is developed under windows operating system. In this program every training function used are implemented strictly to follow the mathematical code of mentioned functions.

The experimental data used for ANN design and training were divided into two separate sets which were used as training and testing data. Considering low number of overall available data due to high cost of experiments for providing them, both sets were used as testing and training. In each time of training one of them were chosen as the training set and the other as the testing set. Following this procedure, all data were affected the overall performance of artificial neural network. Without testing set, artificial neural network may incur overfitting problem. If this phenomena occurs we may add unusual or unnecessary curves in predicted data that will distance us from real behavior of the function and if we use less data for our training we may incur underfitting problem and produced artificial neural network produced won’t be the network we need and its prediction ability will be less than what is expected in the testing set. So selecting a good percentage of data as training and testing will enable us to reach better behavior. Experiments has shown that using 75% of data as training set and the remaining parts as testing set will yield most satisifiable result in data under investigation.

The tested data were normalized before usage in artificial neural network training. As it is shown in table 3 for choosing correct normalization range, different sets of normalized input data were used which were 0 - 1, 0 - 2, and -1 - 1. Also it was tried to use unnormalized data to see whether normalization is beneficial or not. The experiments has shown that if normalization is around zero it will have slightly better performance and the data produced are better mapped to real data.

In this study two different classes of artificial neural networks were used. The first class was consisted of multi-layer feed forward backpropagation network and the other class was recurrent networks where each layer uses its own output as one of the inputs of that layer.
Artificial Neural Networks - Application

<table>
<thead>
<tr>
<th>Normalization method</th>
<th>No normalization</th>
<th>0 - 2</th>
<th>0 - 1</th>
<th>-1 - 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Mean Squared Error</td>
<td>2e-2</td>
<td>4e-4</td>
<td>3e-4</td>
<td>2e-4</td>
</tr>
</tbody>
</table>

Table 3. Effects of normalization on average Mean Squared Error (MSE) in training the recurrent and feed-forward ANN

As recurrent class tries to use more parameters to define the output and input relationship, the vector space defining their relation will have more dimensions. These added dimensions give the network more power than the same network structure without feedback. The base structure used for computation is three layers artificial neural network. This structure is the most used structure in scientific prediction and modeling. The first layer is input layer and the third layer is the output layer and the second layer is the hidden layer which will do the main job of modeling the relationship between input and output data. Each layer has an activation function which was selected from logarithmic sigmoid, tangent sigmoid and linear function. These functions use following formula respectively:

\[
\text{log sig}(n) = \frac{1}{1 + e^{-n}} 
\]

(8)

\[
\text{tan sig}(n) = \frac{2}{1 + e^{-2n}} - 1
\]

(9)

\[
\text{lin}(n) = n
\]

(10)

Each of the three layers will use one of the functions said. If we want to try different combinations of these function 27 different combinations are used in experiments which involves all of the functions.

It can be seen that all of the neurons in the layer will affect the output. Considering this effect we have to determine the right number of neurons in each layer. Input layer will have the number of factors and the output layer is the number of outputs. For determining the number of hidden layer neurons as it is shown in table 4 we have to try different numbers of neurons and deduce the right number by using best result achieved. Different experiments have shown the 13 neurons for feed-forward network and 9 neurons for recurrent network yield the best result.

<table>
<thead>
<tr>
<th>Nodes deleted</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE Recurrent ANN</td>
<td>3e-4</td>
<td>4e-4</td>
<td>1e-3</td>
<td>1e-3</td>
<td>1e-2</td>
<td>2</td>
</tr>
<tr>
<td>MSE Feed-forward ANN</td>
<td>0.5e-3</td>
<td>1e-3</td>
<td>2e-3</td>
<td>2.1e-3</td>
<td>3.6e-3</td>
<td>4e-3</td>
</tr>
</tbody>
</table>

Table 4. Effects of nodes number on average MSE of training ANNs

4. Results and discussion

4.1 Materials characterization

The XRD patterns of synthesized nanoparticles are shown in figure 5. All peaks are attributed to wurtzite ZnO (JCPDS 036-1451) and no other crystalline phases were detected.
The diffractions from (101) plane sets were chosen for estimating the crystallites size of synthesized particles using equation (5).

![XRD Patterns of synthesized ZnO nanoparticles](image)

**Fig. 5.** XRD Patterns of synthesized ZnO nanoparticles

Figure 6 illustrates the UV-Vis spectra of the samples. All samples have an extremum point which is related to generation of exciton in the absorption just below the band edge of the semiconductor. All synthesized particles show a blue shift (Shift to lower wavelengths) in their absorption spectra in comparison with bulk ZnO. The absorption coefficient of the samples was calculated using equation (2). The density of bulk ZnO, nanoparticles concentration and optical path used in this derivation were equal to $5.606 \text{ g/cm}^3$, $3\times10^{-4} \text{ g/cm}^3$, and 1 cm, respectively.

![Absorption spectra of synthesized samples](image)

**Fig. 6.** Absorption spectra of synthesized samples

The exciton band energy was determined by plotting $\frac{da}{d(hv)}$ vs. $hv$, as shown in figure 7.
Fig. 7. Transformed absorption spectra of synthesized samples
The band gap energy of the synthesized nanoparticles were calculated using equation (4) and considering the allowed direct transition probability (n=0.5).

4.2 ANN results

4.2.1 Physical explanation using multi-layer feed forward backpropagation network

The ANN results for samples with 0.75 M initial concentration of Zn(Ac)$_2$ are shown in figure 8. The contours are shown below the 3D surfaces of the modeled properties.

Fig. 8. ANN results for [Zn(Ac)$_2$]=0.75M; a) Crystallite size, b) Band gap energy and c) Exciton energy

As it can be seen in this figure, increasing synthesis time and temperature resulted in bigger crystallite sizes. From thermodynamics point of view, nucleation of new solid phase in aqueous medium has some energy barrier. Temperature can provide the needed energy to
overcome this barrier. It also increases solubility of materials in water which can accelerate the rate of materials solution and precipitation in saturated aqueous medium.

Ostwald ripening is the phenomenon of dissolution of unstable phases and their recrystallization in a more stable form. This process involves solution of materials in water and their precipitation in more stable condition.

Plane defects as grain boundaries and particles surfaces are the main forms of high energy defects in materials which can be partially fixed by crystallites and particles growth. Materials with bigger crystallites have less grain boundary per unit volume so they are more stable. Thus, by increasing synthesis time and temperature, Ostwald ripening processes causes an increase in crystallites size.

When the first stable nuclei of ZnO are formed with sub-nano sizes, due to quantum confinement effects these particles have widest band gap (Wang & Herron, 1991). As crystallite size increases band gap will decrease. When the band gap of particles is equal to bulk material, it can be seen that the band gap energies and the exciton energy bands respectively become lower and higher.

Considering these changes, the overall changes in the properties by increase of time and temperature should be similar to what is shown in figure 8. In conclusion, by increase in time and temperature:

- Crystallite size of the obtained ZnO nanoparticles should increase.
- Band gap energy of the obtained ZnO nanoparticles should decrease.
- Exciton energy of the obtained ZnO nanoparticles should increase.

### 4.2.2 Comparison between recurrent ANN and feed-forward ANN

The ANN models for crystallites size, band-gap and exciton energy of obtained nanoparticles are shown in figures 9, 10 and 11, respectively. As could be seen in these figures, results of models designed by recurrent and feed-forward ANN estimate greatly differ from each other.

![Fig. 9. Crystallite sizes for initial concentration of Zn(Ac)₂ equal to a) 0.5, b) 0.75 and c) 1 molar. Right column results are obtained from recurrent and left column from feed-forward ANN.](image-url)
Fig. 10. Band gap energy for initial concentration of Zn(Ac)$_2$ equal to a) 0.5, b) 0.75 and c) 1 molar. Right column results are obtained from recurrent and left column from feed-forward ANN.

Fig. 11. Exciton energy for initial concentration of Zn(Ac)$_2$ equal to a) 0.5, b) 0.75 and c) 1 molar. Right column results are obtained from recurrent and left column from feed-forward ANN.
As it can be seen in figures 9, 10 and 11, the lack of training data in borders of studied domain causes physically incorrect results around central points. This problem can be solved by getting more experimental results in border area.

As it is shown in table 5, the average MSE achieved in recurrent ANN is much better than feed-forward ANN. Although recurrent network has a better fitness, its results are not physically realistic. This is due to introducing unexplainable curvatures in the parts with insufficient experimental points. This kind of variations in curvature could not be happening from physical point of view due to explanations cited in previous sections. As we have little training data for recurrent network considering it has more parameters than feed-forward one, the recurrent network tries to cope with this problem by adding curves in places where it lacks the data needed and it will try to predict this data and it will underfitted regarding real physical models.

<table>
<thead>
<tr>
<th>Modeled property</th>
<th>Average MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recurrent ANN</td>
</tr>
<tr>
<td>Exciton energy</td>
<td>3e-4</td>
</tr>
<tr>
<td>Band gap energy</td>
<td>2e-4</td>
</tr>
<tr>
<td>Crystallite size</td>
<td>3e-4</td>
</tr>
</tbody>
</table>

Table 5. Comparing the average MSE achieved in recurrent and feed-forward ANN

5. References


This book covers 27 articles in the applications of artificial neural networks (ANN) in various disciplines which includes business, chemical technology, computing, engineering, environmental science, science and nanotechnology. They modeled the ANN with verification in different areas. They demonstrated that the ANN is very useful model and the ANN could be applied in problem solving and machine learning. This book is suitable for all professionals and scientists in understanding how ANN is applied in various areas.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: