A Comparison between Multi-Layer Perceptron and Radial Basis Function Networks in Detecting Humans Based on Object Shape

Laith Jasim Saud
Zainab Kudair Abass
Dept.of Control and System Engineering /University of Technology

Abstract

Human detection represents a main problem of interest when using video based monitoring. In this paper, artificial neural networks, namely multilayer perceptron (MLP) and radial basis function (RBF) are used to detect humans among different objects in a sequence of frames (images) using classification approach. The classification used is based on the shape of the object instead of depending on the contents of the frame. Initially, background subtraction is depended to extract objects of interest from the frame, then statistical and geometric information are obtained from vertical and horizontal projections of the objects that are detected to stand for the shape of the object. Next to this step, two types of neural networks are used to classify the extracted objects. Tests have been performed on a sequence of frames, and the simulation results by MATLAB showed that the RBF neural network gave a better performance compared with the MLP neural network where the RBF model gave a mean squared error (MSE) equals to 2.36811e-18 against MSE equals to 2.6937e-11 achieved by the MLP model. The more important thing observed is that the RBF approach required less time to classify the detected object as human compared to the MLP, where the RBF took approximately 86.2% lesser time to give the decision.

Key words: neural networks, MLP, RBF, object detection, object classification.

المجلد (31) العدد (2) عام 2018   مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Ibn Al-Haitham J. for Pure & Appl. Sci.   *Vol.31 (2) 2018*

Introduction

The problem of detecting humans in images has been extensively considered in literature as they are needed in so many applications in a lot of fields. A main such application is monitoring (surveillance) visually. Monitoring systems can be found in many places like houses, streets, work places, and shops, and this use has been mainly for the purpose of getting records to be used for security purposes or cases alike, and did not involve detecting presence of human objects in images, which is in reality a complex issue by itself and when compared to detecting other types of objects. This is so as humans can assume a variety of postures and are articulated in their shape, so it is impossible to use just one model capable of covering the whole of such possible cases. The first issue a visual monitoring system has to deal with to identify an object is to abstract objects from a given image that are candidates to be compared with a targeted object [1].

To detect an object requires isolating objects of concern in the video frames by clustering pixels of the frame into background and object. Such a thing can be achieved with different methods like "background subtraction", "frame difference", and "optical flow" [1]. The moving part abstracted from the frame (object) could represent unalike moving objects like humans, cars, clouds in sky, birds, and trees that sway. And one way to detect humans among these objects is through classifying these objects. The ways used for object classifications are either based on its shaping, or its movements, or its coloring, or its structure [2]. Classification of the objects is an easy task for humans, but it's challenging to the machine. Object classification includes two stages, candidate objects detection, and pattern recognition which in turn includes two stages which are feature extraction and object classification [3].

Depending on the features extracted, the candidate objects are then classified into pre-specified categories using appropriate methods through comparing the candidate objects pattern with objects patterns in a reference database.

There are different methods to classify the extracted objects like support vector machine, decision Tree, fuzzy classification, and artificial neural networks (ANNs) [3-8].

ANNs have been used and managed to solve such kind of problems that were normally solved by statistical ways. Among such problems they have been used for are, For example, speech recognition, recognition of currents of sonar beneath water, guessing the subaltern construction of round proteins, and classification problems [9]. ANNs are efficient in handling noisy data [10].

There are different types of ANNs. One of the types is the feed-forward ANNs (like the single layer feed-forward nets, multi layer feed-forward nets (MLP), and that called radial basis function (RBF). Another type is the recurrent (feedback) ANNs like the Hopfield networks, Elman networks, and Jordan networks [11].

In this work, the MLP and the RBF are chosen for classification purpose that serves the target of human detection in a sequence of images. Also, an evaluation and a relative comparison will be made among the two methods considering their efficiency in accomplishing the classification. The main reason for choosing MLP is that, with sufficient data, sufficient inner (hidden) units, and suitable time to train, an MLP of one inner (hidden) layer is capable to learn approximation, essentially speaking, of any formula to any level of exactness. That is why, regarding approximation, MLPs are said to be characterized of universality. This means they could be used in case of having little of ahead information that relates the targeted object to inputs. It is true that one inner (hidden) layer is enough in case of having sufficient data, but there are cases in which a net having two or more inner layers might need lesser inner units and weights compared to a net of one inner layer. This leads to the conclusion that additional inner layers occasionally can serve the purpose of generality. The RBF is chosen due to its better approximation capability, simpler network structure, and faster learning algorithms. RBF networks have been widely used in many science and engineering applications. The RBF and MLP are the utmost employed kinds of feed-forward

المجلد (31) العدد (2) عام 2018   مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Ibn Al-Haitham J. for Pure & Appl. Sci.   *Vol.31 (2) 2018*

neural nets. There is a difference between the MLP and the RBF in how the inner units process the data streaming from the inputs. Where, the RBF depends on the Euclidean distance, and the MLP uses inner products. Also, the MLP separates the classes by using hidden units which form hyperplanes in the input space, as indicated in **Fig. 1a**, while RBF separates the classes by local kernel functions, as indicated in **Fig. 1b** [12]**.** Regarding training, most of the methods used for training MLP can also be applied to RBF networks [13]**.**
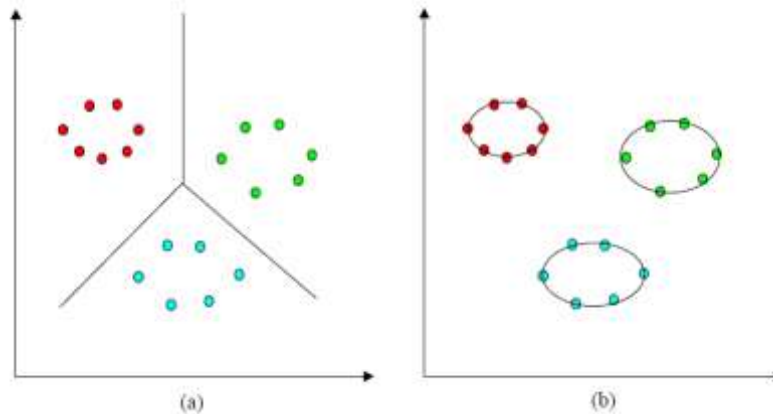

Figure (1): Decision surfaces in two dimensional space: (a) Multi-layer perceptron; (b) Radial basis function network.

The work steps
Candidate Objects Detection

As stated in section 1, the first stage in objects classification is candidate objects detection. To detect candidate objects means isolating objects of concern in video frames. This purpose can be achieved with different approaches like "background subtraction", "frame difference", and "optical flow" [1]. Background subtraction method depends the difference between the current image and the background image to detect objects added to the scene [14]**.** The method formulas are given in Eq. (1) and (2):

$$R_k(x, y) = f_k (x, y) - B (x, y) \tag{1}$$

$$D_k(x, y) = \begin{cases} 1 & back\ ground\ R_k(x,y) > T \\ 0 & target\ R_k(x, y) \qquad \leq T \end{cases} \tag{2}$$

Where f_k is the current frame, B is the background image, and T is the threshold value.
Pattern Recognition

The second stage in candidate objects classification is pattern recognition. The issue of recognizing patterns can be subdivided into two problems, extracting features, and classifying. Next to detection of objects of concern, it is important to abstract some features for recognizing and modeling their shapes in automatic way. Such thing can be achieved through finding a collection of coefficients that provides a substantial description for the information being delivered [1]**.** Regarding classification, the objects are classified based on the extracted features into different categories by using suitable methods that compare the objects of interest with objects inside a reference database, [13]. The approaches used up to now for classifying the objects are either "shaping" based, or "movement" based, or "coloring" based, or "structure" based [2]. The abstracted attributes or features are submitted to the classifying net [1], which (i.e. the classifying net) is an artificial neural network in this work.

المجلد (31) العدد (2) عام 2018   مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Ibn Al-Haitham J. for Pure & Appl. Sci.   *Vol.31 (2) 2018*

To get some features about the object, some steps need to be done. Firstly, the system produces a different form for the binary image of the targeted object. Then, based on this newly produced form, the most important statistic and geometric properties are extracted. This new representation is composed of two projections, vertical and horizontal. Where, the vertical projection is given by the sum of the white points in the rows of the binary image, and the horizontal projection is given by the sum of the white points in the columns of the binary image and as given by the following formulas [1].

$$H\,pro(j) = \sum_{i=1}^{N} I(i.j) \tag{3}$$

$$V\,pro(i) = \sum_{j=1}^{M} I(i.j) \tag{4}$$

Where, M represents columns' number in the binary representation of the image I and N is the rows' number.

From Eq. (3) and (4), seven parameters (PR1-PR7) are evaluated which represent the object shape. The parameters are:

1. Vertical projection largest value:
$$PR1 = \max imum\,\{V\,pro(i)\} \tag{5}$$
2. Horizontal projection largest value:
$$PR2 = \text{maximum}\{H\,pro(j)\,\} \tag{6}$$
3. "Vertical projection" factors summation (similar for "horizontal projection"):
$$PR3 = \sum_{i=1}^{N} V\,pro(i) = \sum_{J=1}^{M} H\,pro(j) \tag{7}$$
4. "Vertical projection" mean value:
$$PR4 = \sum_{i=1}^{N} V\,pro(i)\,/\,H \tag{8}$$
Where; H represents the quantity of components of nonzero value in the "vertical projection" (N ≥ H).
5. "Horizontal projection" mean value:
$$PR5 = \sum_{j=1}^{M} H\,pro(j)/k \tag{9}$$
Where; K represents the quantity of components of nonzero value in the "vertical projection" (M ≥ K).
6. "Vertical projection" standard deviation:
$$PR6 = \sum_{i=1}^{N} |V\,pro(j) - \sum_{i=1}^{N} V\,pro(i)\,/H| \tag{10}$$
7. "Horizontal projection" standard deviation:
$$PR7 = \sum_{j=1}^{M} |H\,pro(j) - \sum_{j=1}^{M} H\,pro(j)/k| \tag{11}$$

Artificial neural networks

In recent decades, neural computing has emerged as a practical technology for the purposes of classification, function approximation, data processing, filtering, clustering, compression, decision making, etc., with successful applications in many fields as diverse as medicine, finance, geology, engineering, biology and physics [9 &11].

An artificial neural network can be defined as a set of simple computational units that are highly interconnected. The units are loosely representing the biological neurons   and are also called nodes. Figure (2) gives a depiction for the neuron [1].
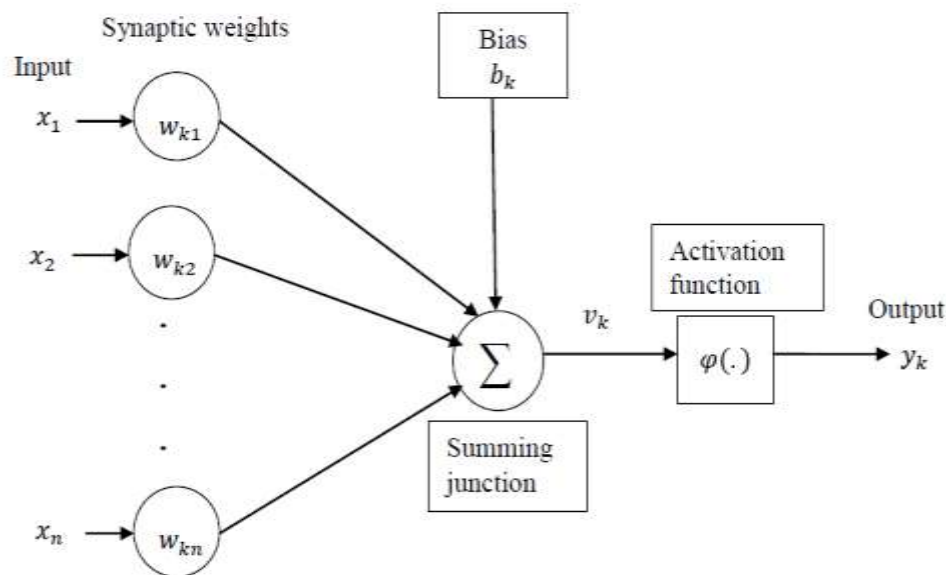
المجلد (31) العدد (2) عام 2018

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Ibn Al-Haitham J. for Pure & Appl. Sci.

Vol.31 (2) 2018

Figure (2): Graphical presentation of neurons.

A neuron can be defined as an information processing unit that is fundamental to the operation of a neural network. The connections between nodes are unidirectional. Such connection system resembles those "synaptic connections" of a brain. For each individual connection there is a weight, (w_{kj}), called "synaptic weight" and stands for how the connection is strong between units j and k. This weight may take positive or negative value which makes it dissimilar to brain's synapse. For a given node (neuron), the outputs of the nodes connected to it are summed, after being multiplied by specified weights, to form its input. Then, this input will be modified by what is known as the "activation function" of the node. The "activation function" is also known as "squashing function". In reality this function modifies the input value within limits. The model of such a node (neuron), which is shown in figure (2), contains a "bias", symbolled by b_k, that can lower (if it is positive) or increase (if it is negative) the net input of the "activation function" [1].

In mathematical terms, a neuron *k* may be described by the following equation [1]:

$$y_k = \varphi(v_k) = \varphi(\sum_{j=1}^{n} w_{kj} x_j + b_k) \tag{12}$$

Where;

x_1 to x_n represents the "patterns" of the input.

w_{k1} to w_{kn} represents the "synaptic weights" for node (neuron) k.

b_k stands for the "bias".

$\varphi(\cdot)$ represents the "activation function".

y_k represents the neuron output.

And the RBF neural network formula is [15]:

$$y = \sum_{i=1}^{N} w_{ij} \varphi(x_k, x_i) \tag{13}$$

When the activation function is Gaussian function:

$$\varphi(x_k, x_i) = \exp\left(-\frac{1}{2\sigma_i^2} \sum_{m=1}^{M} (x_{km} - x_{im})^2\right) \tag{14}$$

The types used in this work, as indicated and justified in the introduction, are the MLP and the RBF.

المجلد (31) العدد (2) عام 2018   مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Ibn Al-Haitham J. for Pure & Appl. Sci.   *Vol.31 (2) 2018*

Experimental Results

Datasets from Weizmann database walk action (low-resolution (180x144) video clips with 9 different subjects/10 videos) running at 25 frames per second have been used in the experimentations in this paper.

The work included three main steps:

Step1: Extract the objects of interest from an image by using background subtraction equation. The threshold value used in this work equal to 30 and some of result obtained from this stage can be showed in figure (3). This stage succeeded in detect the object in 10 video out of 10, this mean the ratio of succeed in this stage is 100%.

Step2: Extract some features that represent the object shape by using shape based method. In this work, 7 parameters (p1 to p7) are extracted which represent the object shape and then used as input to the classifier which is a feedforward neural network.

Step3: Using two types of artificial neural networks (MLP and RBF) to classify the extracted objects into human or non-human. The Neural networks model implemented by using neural networks toolbox function in mat-lab which is (newff) for multilayer perceptron and (newrb) for radial basis function, these tools has the ability to generate initial weights and biases for the network and save the values of weights, biases from training to used it in testing. MLP in this work consists of 2 hidden layers. A sigmoid transfer function is used in the two hidden layers in order to get a true or false output. The model used here is shown in figure (4). RBF has a standard structure which consists of one hidden layer with Gaussian activation function and one output layer with linear activation function. Some of the parameters being set up in the RBF network to get good performance are the spread and the maximum number of neurons in the hidden layer. The model used for this case is shown in figure (5). The samples of images used to train the two NNs models are given in figure (6). The samples of images used to test the two NNs models are given in Fig. 7. And some classifiers results can be showed in figure (8).
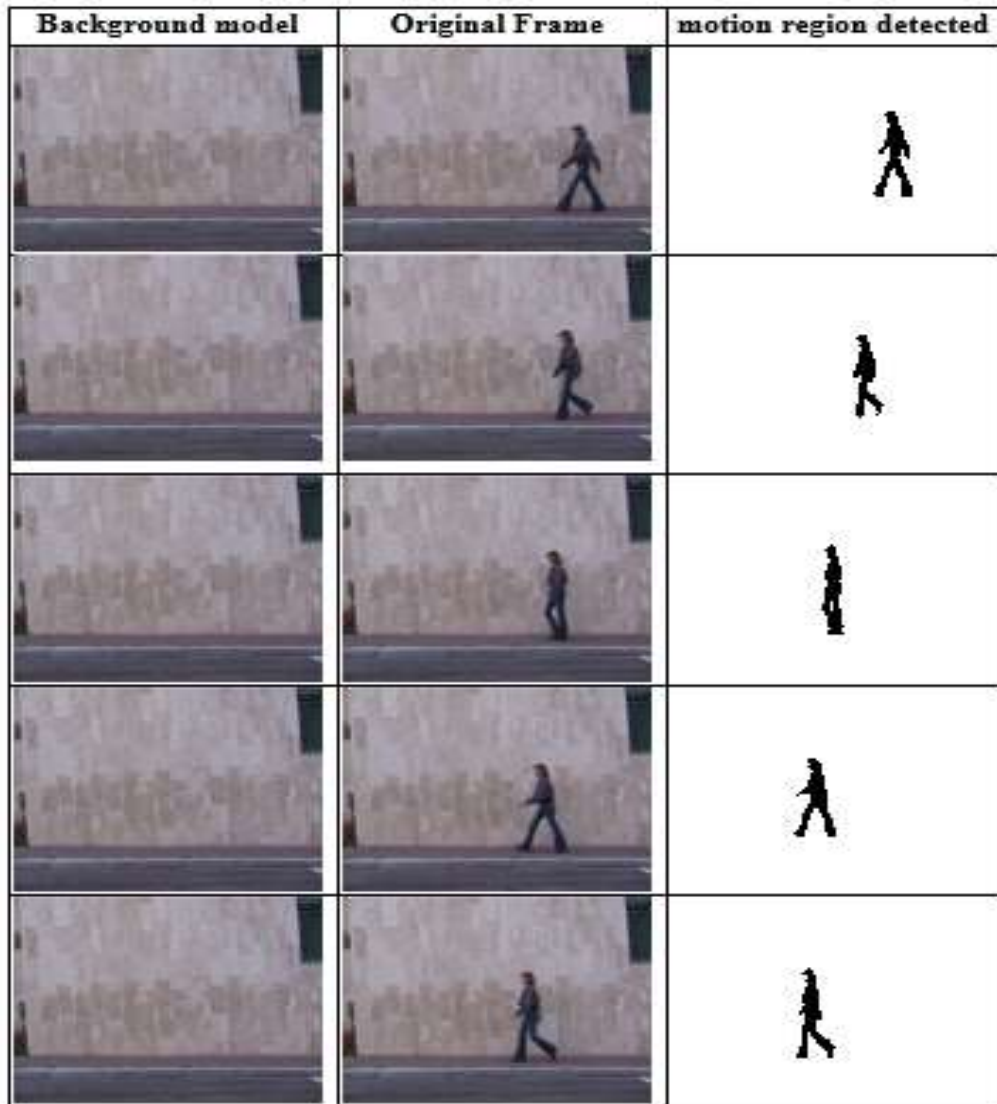
المجلد (31) العدد (2) عام 2018

Ibn Al-Haitham J. for Pure & Appl. Sci.

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Vol.31 (2) 2018

Figure (3): **Background model, Original frames and background subtracted frames with motion region detected.**
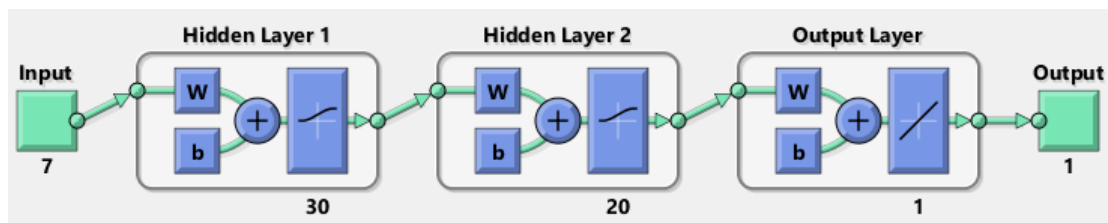
**Figure (4): The multilayer perceptron net used in the experimentations.
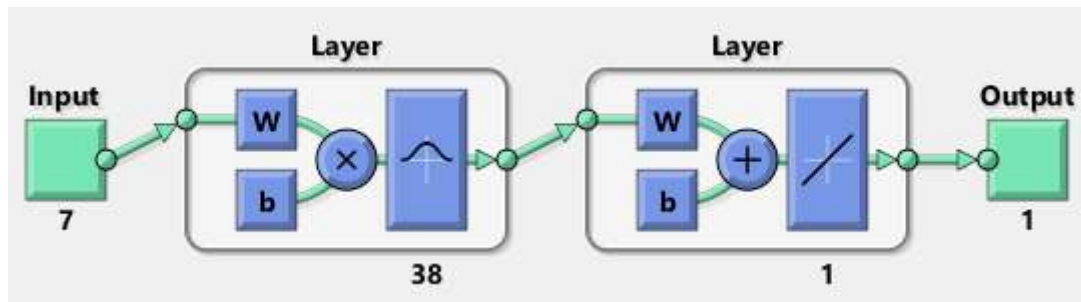Where: w = weight, b = bias.**

المجلد (31) العدد (2) عام 2018

Ibn Al-Haitham J. for Pure & Appl. Sci.

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Vol.31 (2) 2018

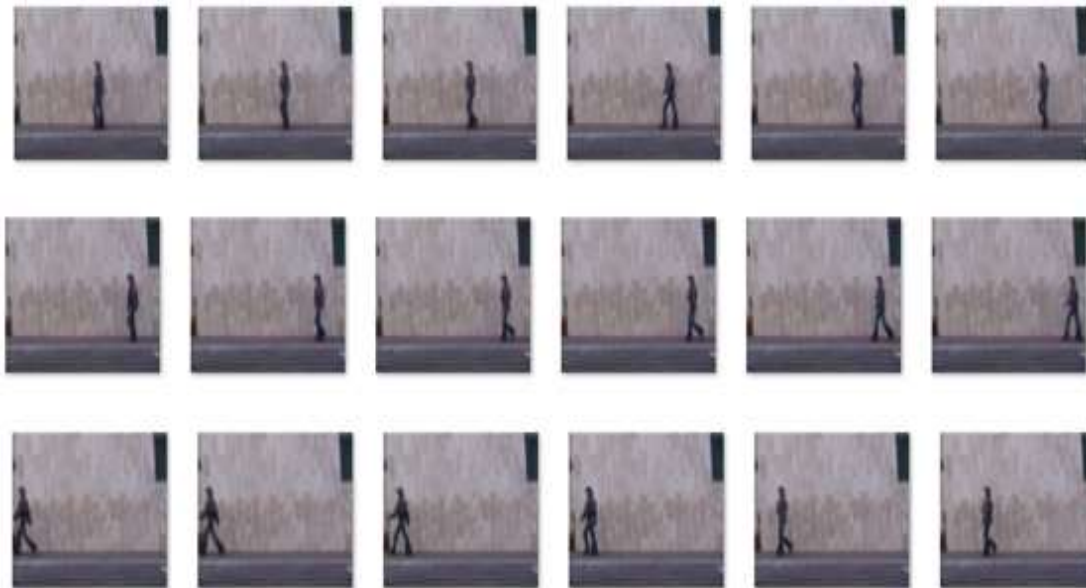**Figure (5): The radial basis function net used in the experimentations**.

Figure (6): The images samples used to train the NNs models

.

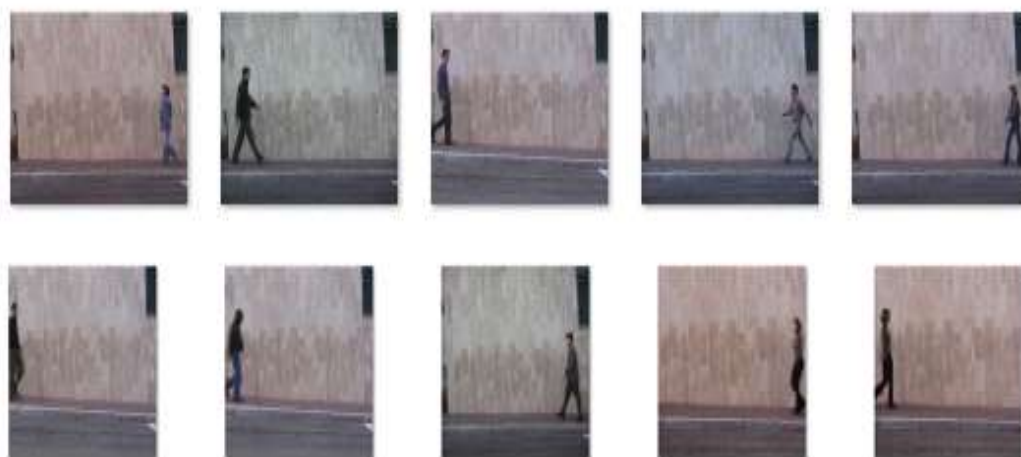**Figure (7): The images samples used to test the NNs models**

المجلد (31) العدد (2) عام 2018

Ibn Al-Haitham J. for Pure & Appl. Sci.

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Vol.31 (2) 2018

Difference image	Result of classifier (human ?)
	Yes
	Yes
	Yes
	Yes
	Yes

Figure (8): Sample of results obtained from Classifiers

The computer codes for the MLP and the RBF models were developed in Mat-Lab software (version 2016). The models were trained until the best performance was obtained. The optimal parameters (biases and weights) of the network were saved and used for testing and validating operation of the models that are given in figure (4 and 5), with their properties that are given in table (1). This stage succeeded in classify the object in 10 video out of 10, this mean the ratio of succeed in this stage is 100%.

Table (1): Optimum properties of the ANNs

Parameter	MLP	RBF
No. of layers in the network	4	3
No. of nodes in input layer	7	7
No. of Hidden layers	2	1
No. of nodes in Hidden layer	1st layer=30 2nd layer=20	38
No. of nodes in output layer	1	1
Activation function of hidden layer	1st layer: sigmoid 2nd layer: sigmoid	Gaussian
Activation function of output layer	Linear	Linear
No. of iterations	575	Non
Number of neurons to add between displays	Non	1
Mean squared error goal	0	0
Spread	Non	1e+11
Maximum number of neurons	Non	50

The results obtained are shown in figure (9) which explain that the error reduces after more epochs of training and the best performance is taken from the epoch with the lowest

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية | المجلد (31) العدد (2) عام 2018

Ibn Al-Haitham J. for Pure & Appl. Sci. | *Vol.31 (2) 2018*

validation error and Fig. 10 which explain that the error reduces after more epochs of training until it reaches to minimum error. where MSE is the mean squared error and it represents an estimator's quality measuring. This value is at all times either zero or positive. Values near to zero represents a good indication of quality, and are summarized in table (2) for performance and Table (3) for time.
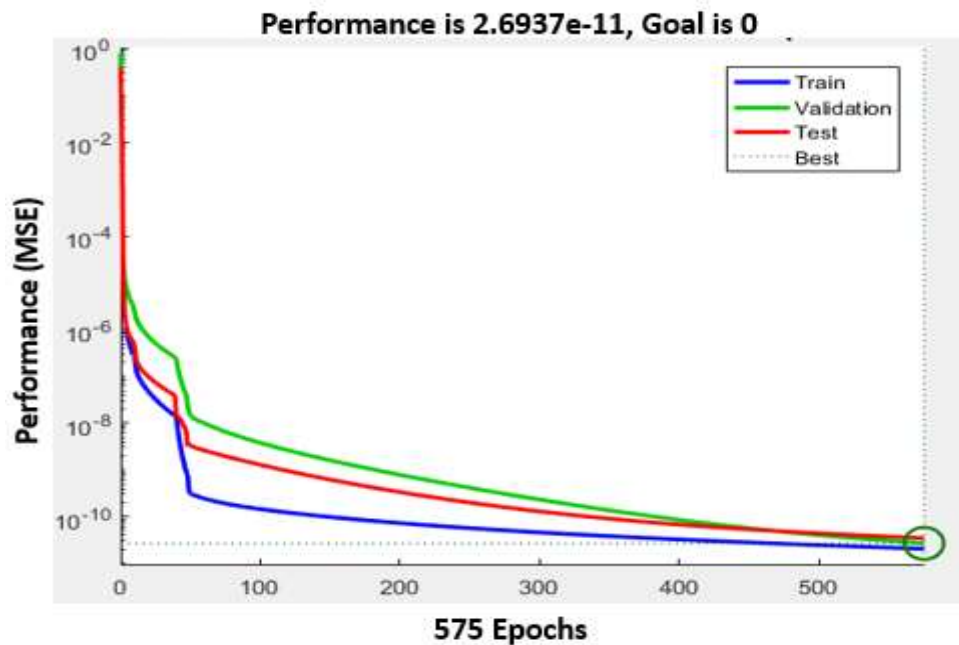

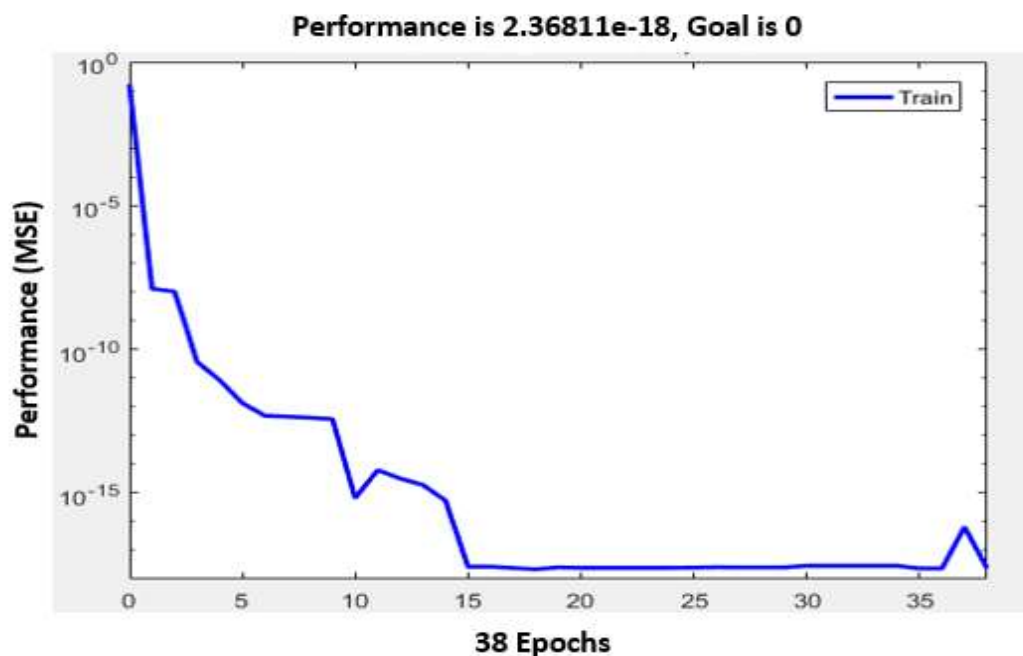**Figure (9):** Performance of the MLP neural network.

Figure (10): Performance of the RBF neural networks

المجلد (31) العدد (2) عام 2018

Ibn Al-Haitham J. for Pure & Appl. Sci.

مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Vol.31 (2) 2018

Table (2): Performance indices for the tested nets models.

Model	MLP	RBF
Performance	2.6937e-11	2.36811e-18

Table (3): Time indices for the tested nets models.

Model	MLP	RBF
Profile time	51 sec.	7 sec.

Conclusions

From simulation results it is found that the classification performance of both of the RBF and the MLP types of neural networks is acceptable. Still, a better relative performance was achieved by the RBF model considering the MSE as it gave a MSE of 2.36811e-18 at epoch 38 compared to a MSE of 2.6937e-11 at epoch 575 given by the MLP, the error goal of both models was zero which can be seen in ƒigure (9) which explain that the error reduces after more epochs of training and the best performance is taken from the epoch with the lowest validation error and Fig. 10 which explain that the error reduces after more epochs of training until it reaches to minimum error. where MSE is the mean squared error and it represents an estimator's quality measuring. This value is at all times either zero or positive. Values near to zero represents a good indication of quality. Also it is noted, and this a crucial issue, that the RBF took less time than the MLP, where the RBF took 7 seconds and the MLP took 51 seconds to give the output (i.e., to decide if the detected object is human or not), and that is because for model development in the RBF case, no repetition is required to reach the optimum model parameters as is the case with the MLP. This represents approximately a 86.2% lesser time.

References

1. Leo, M; Spagnolo,P; Attolico, G and Distante, A. (2003) Shape Based People Detection for Visual Surveillance Systems,.
2. Parekh, H. S; Thakore,D. G and Jaliya,U.K. (2014) A Survey on Object Detection and Tracking Methods, International Journal of Innovative Research in Computer,.
3. Dhaware, C. K; Wanjale,H. and Survey (2016) On Image Classification Methods in Image Processing, International Journal of Computer Science Trends and Technology,.
4. Bianchini,M and Scarselli, F.jk,x 2014 On the Complexity of Neural Network Classifiers: A Comparison Between Shallow and Deep Architectures, IEEE Transactions on Neural Networks and Learning Sysyems,.
 5. Kamavisdar,P; Saluja,S; Saluja, and Agrawal,S; (2013) A Survey on Image Classification Approaches and Techniques International Journal of Advanced Research in Computer and Communication Engineering,.
 6. Darlin, R; Nisha, R; Gowri,V and Scholar,PG (2014) A Survey on Image Classification Methods and Techniques for Improving Accuracy, International Journal of Advanced Research in Computer Engineering and Technology.
 7. Naswale, P; Ajmire, E and Image Classification Techniques-(2016) A Survey International Journal of Emerging Trends and Technology in Computer Science.

المجلد (31) العدد (2) عام 2018   مجلة إبن الهيثم للعلوم الصرفة و التطبيقية

Ibn Al-Haitham J. for Pure & Appl. Sci.   *Vol.31 (2) 2018*

8. Siddhartha,S; Girish, N; Jajnyaseni,M and Sayan, K; (2014) C. A Survey of Image Classification Methods and Techniques, International Conference on Control, Instrumentation, Communication and Computational Technologies,.

9. Jha, G and Artificial K. Neural (2010). Networks, Indian Agricultural Research Institute Library Avenue, New Delhi,

10. Wang,Fu and Wang,Fe; and (2014) Rapidly Void Detection in TSVs with 2-D X-Ray Imaging and Artificial Neural Networks, IEEE Transactions on Semiconductor Manufacturing,.

11. Krenker, A and Bešter, J. A. (2011) Kos, Introduction to the Artificial Neural Networks, Artificial Neural Networks - Methodological Advances and Biomedical Applications,

12. Xie, T and Wilamowski, H. B;(2011). Comparison between Traditional Neural Networks and Radial Basis Function Networks,

13. Sereno, F; Marques, J. P;Matos, A and Bernardes, J. A (2000) Comparative Study of MLP and RBF Neural Nets in the Estimation of the Foetal Weight and Length,

14. Alex,S and Wahi, A; (2014). BSFD: Background Subtraction Frame Difference Algorithm for Moving Object Detection and Extraction, Journal of Theoretical and Applied Information Technology.

15. Xi-mei, L; Xiao-hui, Y; Qian, Z; and Hong-mi, G; (2012) Application of RBF Neural Network in Fault Diagnosis for Transmission Gear, Advanced Materials Research,, 7563-7568.