ARTIFICIAL INTELLIGENCE IN MONITORING SYSTEM

Lucjan Pelc, Artur Smaroń, Justyna Stasieńko

***Abstract**: The article presents the neural network constructed in order to use it for monitoring. Its role is to recognize the events in alarm situations (theft, burglary etc.). The film presenting a real break-in into the car was used while testing this network. The main task of monitoring system based on the neural network is to compile such a network which shows the alarm situations as soon as possible using the available equipment.*

***Keywords**: neutral network, monitoring system, neuron.*

***ACM Classification Keywords**: I.2 Artificial Intelligence – I.2.6. Learning*

Introduction

The monitoring programs are often used in order to gather the information about the amount and the quality of the observed object. The gathered information makes it easier to make right decisions, especially if this state is hazardous for human being and the surrounding. It allows also to improve or remove the results in the existing situation. The article aims at showing one of these solutions. There are many monitoring systems available on his market. They are different as far as the quality, functionality and price are concerned. If we take into consideration the functionality the following types can be distinguished: systems which record on the continuous basis, systems which record if the movement is detected and those which record before and after the incident. The current monitoring systems are able to: record on the continuous basis, detect the movement and record the image, detect the disappearance of the image, follow and count the objects, control the cameras, one or more visual channels, and inform the operator by an e-mail or sms. The cost of such systems amounts to 300 – 10000zl. Unfortunately, there are not many systems which can define the situation before the burglary or theft etc.

The majority of systems posses an insignificant or do not posses any functionality as far as the process of supporting the system operator is concerned especially in case of recognizing important facts and events. The monitoring systems are used rather for recording and gathering the courses of events, even undesirable ones. It requires large disk storage for recording an image. In some systems (e.g. Taiwanese ACTI- APP-2000-32) an intelligent management of memory was used by means of recording the events before they occur which is often called “buffering the image". If the system posses such ability, it buffers the image all the time and in the moment of detection the alarm it has also the images recorded before the accident, for example 5sec. Supporting the system operator while recognizing undesirable situations on the recorded image and responding automatically requires from the system to solve problems connected for example with pattern recognition. One of the authors of this article used probabilistic timed automata which if connected with particular spheres on the observed image allowed to define important actions in reference to the given problem [Pelc, 2008]. This article is a kind of a supplement and expansion of that approach. Instead of using probabilistic timed automata the neural network was used here.

While creating the monitoring system three rules should be taken into consideration: the periodicity of measurement, the unification of the equipment and methodology used for the measurement and observation as well as the unification of the results interpretation. This case study uses a film from the Internet showing the real break-in into the car [http://www.youtube.com/watch?v=pLKjm2uGrU4&feature=related]. It resulted in the idea of creating the monitoring system based on one of the Artificial Intelligence applications "neural network"

The aim of article

The aim of this article is to create a low-cost monitoring system which can be compiled on the basis of computers designed for the general use as well as typical cameras (e.g. network cameras)

The system assumptions:

- one or two cameras
- limited computability of the equipment and disk
- better functionality than in typical monitoring systems
- algorithm of artificial intelligence – the neural network whose computability is not very complex
- the ability to recognize particular actions and reactions depending on their character:
  - ✓ neutral actions – normal work of the system
  - ✓ suspicious actions – generating the warnings by the system, starting recording the sequence of frames
  - ✓ prohibited actions – alarming, calling the operator and recording the sequence of frames

The main feature which distinguishes the monitoring system from many others is its ability to take decisions in order to prevent the prohibited actions and not only to record them. The following method of solving the problem was accepted:

- action was defined as the occurrence of the given movement trajectory in particular period of time
- the movement trajectory is determined on the basis of recognition of the object being observed in the successive spheres of observation
- resignation of the constant recording of the images in aid of the recording which is caused by events
- the event causing the recording is the occurrence of the beginning of the action
- introducing the spheres of observation on the image recorded by the camera limiting at the same time the size of the in-put data.

The method of defining the spheres of observation

The image coming from the camera is covered by the network of spheres of observation consisting of 125 elements. The size of the network is determined automatically depending on the picture definition. The network density can change adjusting the image from the camera. The network is related to the scene or objects being observed whose position may change. However, the main stress is not put here on fluent change of the position and following the object.

The example of using the network related to the scene can be the observation of the gateway. However, in case of the network connected with the object, it can be presented by means of observation of a car parked in different ways and positions. To sum up, the steps of defining the spheres are as follows:

(1) determining if the spheres should take the form of a uniform or condensed network
(2) choosing the object or the scene
(3) indicating the essential points of the image for which the system should condense the network
(4) if needed, defining the size of grid's fields which is different than the implicit one (implicitly there are three types of the size and density of grid's fields: large, medium and small)

(5) in case of the network related to the particular object there is a need of creating the definition of this object. The system defines the object on the basis of its simplified shape, colour or its texture.

(6) If the system is going to be used in applications whose shape does not change or the changes are small than in (5) the size of the object is also taken into consideration. It is worth remembering that in the systems with cameras there is an apparent change of the object size depending on its distance from the camera.

Fig.1 shows the example of the image seen by means of a camera with a network characterized by reduced density. The biggest density is intended for the observed vehicles and it is the situation when the network is related to the observed objects.

Fig 1. An example of the definition of the sphere of observation. Source: Own elaboration.

The logic of the created system

As it has been mentioned earlier, the system should inform about the potential risk of the occurrence of the prohibited action, which as a result, allows to take remedial. In consequence, the logic of the system should provide the proper sensitivity as far as the recognition of the action is concerned and it should also be characterized by the ability of limited prediction. It should be remembered that one of the assumptions at the beginning of this article was the system's ability to recognize the neutral, suspicious and prohibited actions. Detection of the prohibited action is connected with alarming and calling the operator and that is why the system cannot do it precipitously. In other words, the system cannot be too sensitive and its reactions should not be exaggerated. It is especially important in case of determining the prediction. Taking into consideration all of these assumptions, the logic of the system should demand the occurrence of one of these three sorts of action on the basis of the information in the defined spheres of observation. The classic neural networks were used to support this solution. It was necessary to decide about the number of layers in the network. It was assumed that the exit layer will consist of at least three neurons indicating the occurrence of a particular kind of action. In the entrance layer, there should be as many neutrons as the spheres of observation, however, this solution would be enough only in case of recognizing the statistical actions. In the presented example the time lag also should be mentioned. Time is measured with the occurrence of successive frames. That is why, it is necessary to add one more neuron informing about the number of frames which have just appeared. It allows for taking into consideration their movement dynamics, and not only the statistical trajectory. Therefore, it can be assumed that in the "minimal" version the number of neurons in the entrance layer amounts to the number of spheres of observation enlarged by one.

Classic neural network consists of one or few hidden layers. In the experiment the number of the hidden layers and the number of the neurons in each of them were the object of the simulation analysis whose results will be soon discussed in this article. Taking into consideration the previous assumptions referring to computability, two and more hidden layers were analyzed.

The simulation

The aim of the simulation was to determine the construction of neural network which is the most suitable from the viewpoint of the given problem. Such a network should for example:

- recognize defined actions accurately
- predict the occurrence probability almost without any mistakes
- include the least neurons
- learn quickly

According to the assumptions accepted previously the network with one and two hidden layers were analyzed paying particular attention to the accuracy of recognizing the actions as well as fast learning.

Fig.2, Fig.3, and Fig.4 present hypothetical and simplified movement trajectories referring to three types of actions discussed previously. Fig.2 deals with the forbidden action, Fig.3 is related to the suspicious action and Fig.4 presents the neutral action. The protected sphere is marked with "x" and darkened. The black squares represent the recorded trajectory of movement. The number under each of the pictures informs about the number of frames falling on the particular trajectory.

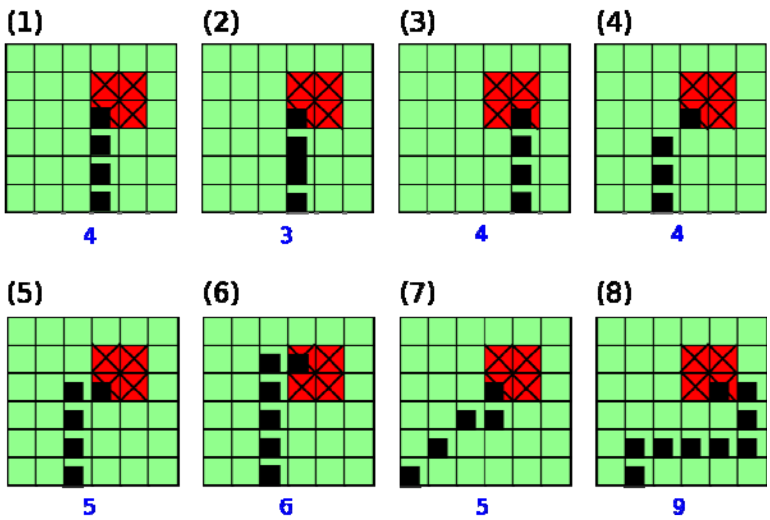

Fig. 2. Examples of movement trajectory for the forbidden-alarm action. Source: Own elaboration.

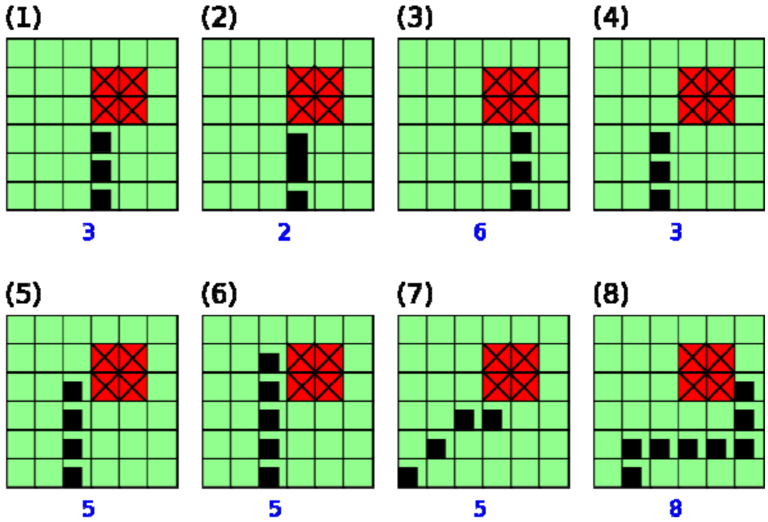

Fig. 3. Examples of movement trajectory for the suspicious-warning action. Source: Own elaboration.

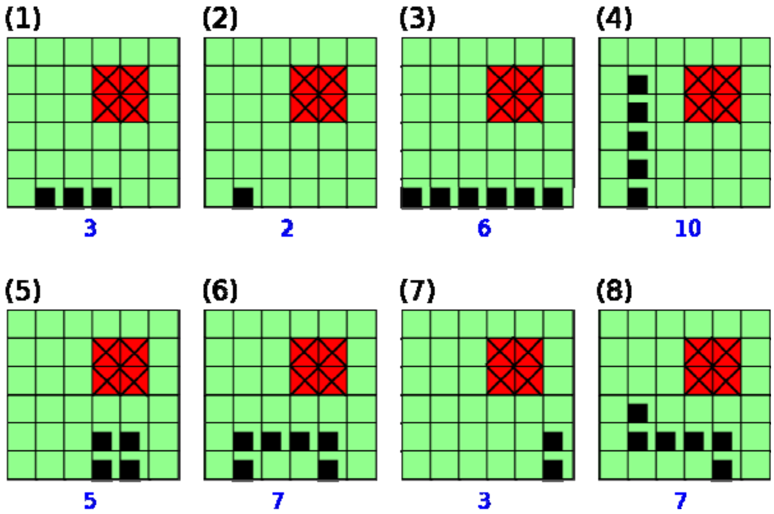

Fig.4. Examples of movement trajectory for the neutral-normal action. Source: Own elaboration.

The elementary picture (1) presented by Fig.2 represents the trajectory going in four steps from the bottom part directly to the marked sphere. Below there are some situations, which could occur for the discussed example of the scene:

- The prohibited action – if the trajectory ends or goes through the marked sphere (Fig.2)
- The prohibited action which is very probable – if the trajectory with great dynamics goes towards the marked sphere (compare the elements (1) and (3) from Fig.2)
- The suspicious action – when the trajectory goes towards the marked sphere but does not reach it, the movement dynamics is not important here
- The suspicious action which is very probable - when the trajectory is near the marked sphere regardless of the dynamics (compare elements (5), (6), (7), (8) from Fig.3)
- The neutral action – the movement trajectory is in a short distance away from the marked sphere regardless the dynamics (Fig.4).

Even after analysing Fig.2 roughly it can be seen that if the object moves quickly on the trajectory with the marked sphere, the system should inform as quickly as possible the occurrence of the prohibited action before the trajectory reaches the marked sphere. In practice, it would allow to prevent the prohibited action to reach the marked sphere in the given example. The permanent persistence of the suspicious action may indicate the potential occurrence of the prohibited action. The situations presented on Fig.3 and Fig.4 were chosen in such a way that the situations which at the beginning looked like forbidden after the occurrence of the successive frames turned out to be suspicious for example Fig.3 (1) and Fig.2 (1), Fig.3 (2) and Fig.2 (2) etc. Taking into consideration the assumptions that the system should be characterized by the ability to predict the situation, but it should not react precipitously, such kind of situations are interesting.

The simulation results

During the simulation it was checked how quickly the network with one or two hidden layers is able to learn as well as the accuracy of recognition if the action is forbidden, suspicious or neutral.

- **The comparison of the capacity of learning for the network with one or two hidden layers**

Below, there are the tests results of the capacity of learning for the network with one hidden layer depending on the number of the neurons in the hidden layer. Fig.5 presents the reduction of mistakes in the process of learning depending on the number of epochs. Other progresses refer to various numbers of neurons in the hidden layer. It is important to emphasize that at the beginning the number of neurons in the hidden layer improves the parameters of learning but after exceeding the certain number of the neurons the situation changes the other way round. On the basis of the conducted analysis the formula was defined. It counts the number of neurons in the single hidden layer as a quotient of the product and the sum of the number of neurons in the input and output layer in the network (1).

$$N_h \approx \frac{N_{in} * N_{out}}{N_{in} + N_{out}} \tag{1}$$

where

N_{in} – the number of neurons in the entrance layer

N_{out} – the number of neurons in the exit layer

N_h – the number of neurons in the hidden layer.

If the number of neutrons in the entrance layer is much bigger than the number of neurons in the exit layer (at least of an order of magnitude), the radical dependence, which combines the number of neurons of the hidden layer with these from the entrance and exit layer (2), seems to be more accurate.

$$N_h \approx 1.2 * \sqrt{N_{in} * N_{out}} \tag{2}$$

The analogous simulations were carried out with two hidden layers. The number of the neurons in the first hidden layer is chosen on the basis of relations resulting from formula (2).

The comparison of progresses for various number of neurons in particular hidden layers presented by means of Fig.6 indicates the validity of accepting the correlation that the number of neurons in the second layer should be smaller by half than the number of neurons in the first layer (according to the rule of pyramids) (3).

$$N_{h2} \approx 0.5 * N_{h1} \tag{3}$$

Having analysed the networks which learned faster with one or two hidden layers it was included that the network with only one layer is able to learn faster.

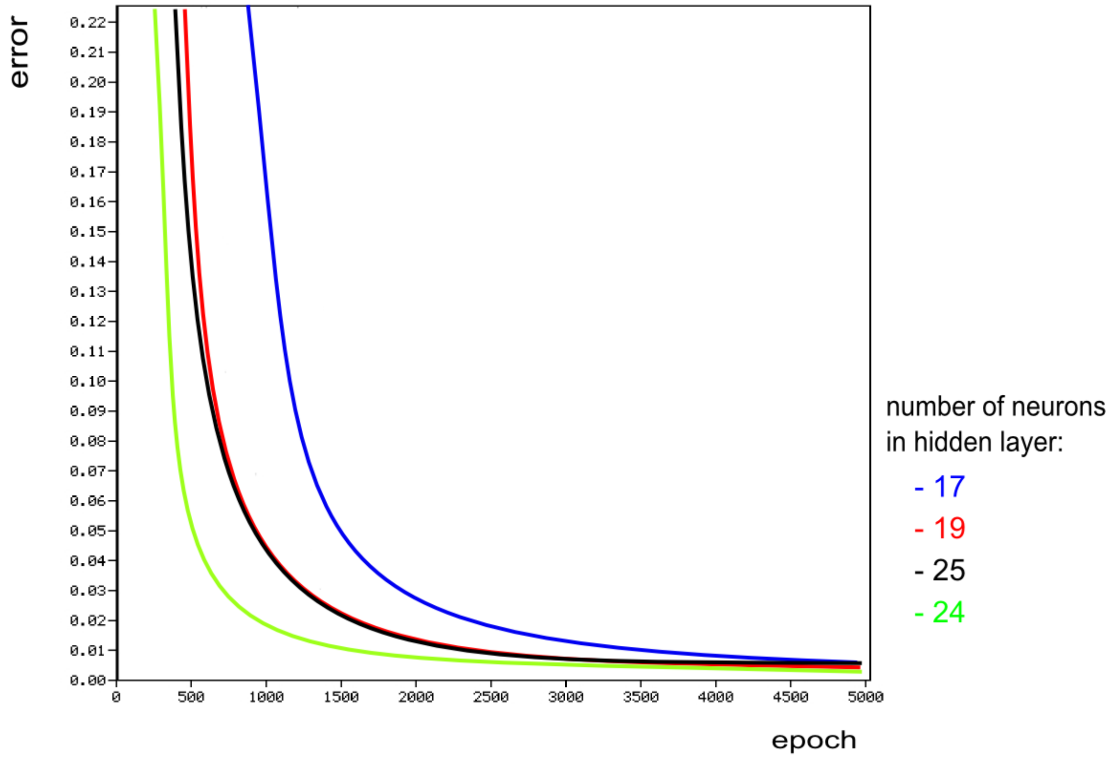

Fig.5. The capacity of learning depending on the number of epochs. Source: Own elaboration

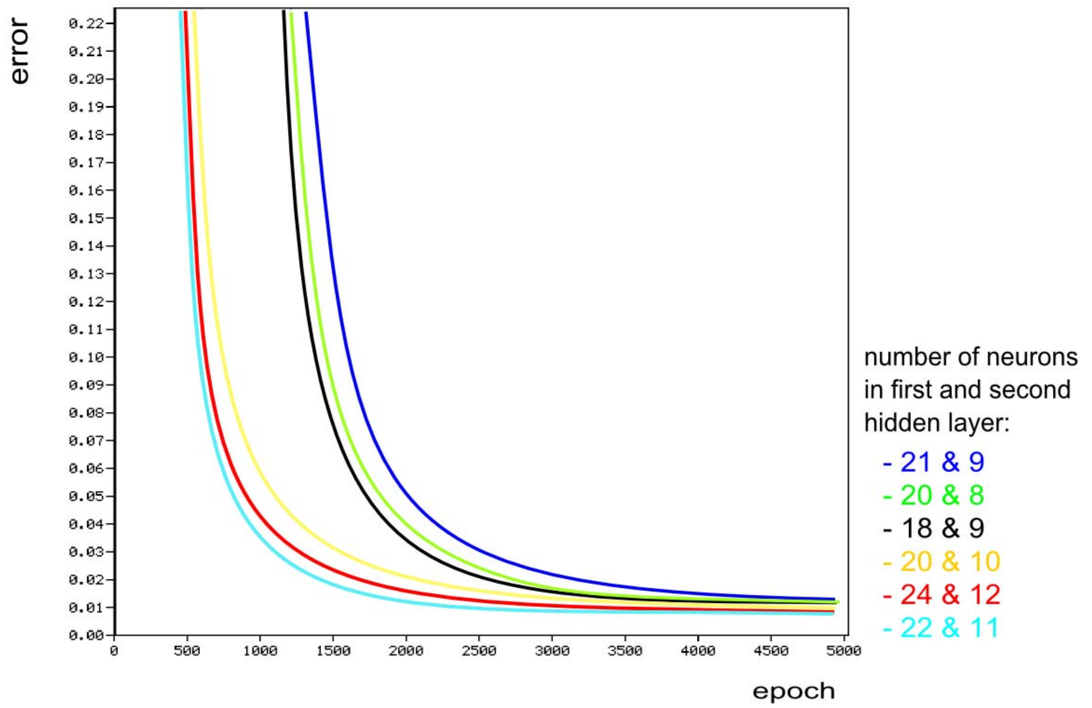

Fig.6. The capacity of learning depending on the number of neurons for two hidden layers. Source: Own elaboration.

- **The comparison of the accuracy in actions recognition**

It turned out that the accuracy of the recognition of the action is bigger for the network with two hidden layers than the network with just one hidden layer. However, in this case the recognition of the forbidden action during the trajectory analysis was precipitate. In consequence, it was the trajectory mostly connected with the suspicious action. That is why, at the dynamic recognition of the action or rather the prediction of the action, the network with one layer turned out to be more accurate solution. It is illustrated by Table 1 which shows how the neuron network recognized the forbidden-alarm action, and suspicious-warning action in the particular period of time, described in Table as steps. Analyzing for example case, it is easy to notice that the network with two hidden layers reacts too rapidly and notices in the developing warning situation the alarm situation. It can be easily seen in steps 4 and 5 where the network guesses the alarm situation wrongly with the probability bigger than 99%. In the successive steps the network recognizes the warning correctly, but taking into consideration the previous assumptions such behavior of the network is delayed. For the same steps (4 and 5) the network with one layer recognized the alarm with the probability of 1,5% and the proper warning situation with more than 92%. Although the network with one hidden layer recognizes the finished actions less accurately than the network with two hidden layers, for the investigated application the network with one hidden layer is more suitable.

Table 1 Example situations recorded by monitoring system

Example situations recorded by monitoring system		Response of the system in percentage					
		One hidden layer – 19 neurons			Two hidden layers – 22 and 11 neurons		
		alarm	warning	neutral	alarm	warning	neutral
Case 1: warning	Step 1	1,72%	0,19%	98,09%	28,06%	0,08%	71,86%
	Step 2	0,82%	0,41%	98,76%	22,72%	0,08%	77,20%
	Step 3	2,85%	1,93%	95,22%	39,02%	0,08%	60,90%
	Step 4	1,55%	92,43%	6,02%	99,13%	0,79%	0,08%
	Step 5	1,49%	94,62%	3,89%	99,12%	0,80%	0,08%
	Step 6	0,67%	98,83%	0,50%	1,77%	98,23%	0,00%
	Step 7	0,60%	99,03%	0,37%	0,58%	99,42%	0,00%
	Step 8	0,55%	99,08%	0,36%	0,48%	99,52%	0,00%
	Step 9	0,73%	98,97%	0,30%	0,34%	99,66%	0,00%
Case 2: alarm	Step 1	6,73%	0,00%	93,26%	5,16%	0,09%	94,76%
	Step 2	1,61%	0,24%	98,15%	56,15%	0,08%	43,77%
	Step 3	0,87%	0,95%	98,18%	73,90%	0,10%	26,00%
	Step 4	1,44%	7,81%	90,74%	94,39%	0,15%	5,46%
	Step 5	97,20%	0,64%	2,16%	98,70%	0,26%	1,04%
	Step 6	76,49%	23,40%	0,10%	2,76%	97,24%	0,00%
	Step 7	98,65%	1,28%	0,07%	12,84%	87,16%	0,00%

Source: Own elaboration.

The applied algorithm of learning

In the elaborated program was used the algorithm of error backpropagation based on the example found in the literature (4).

$$\delta_{i,j} = O_{i,j}*(A_j - O_{i,j})*(1 - O_{i,j})$$

$$\delta^{t}_{i,j} = \delta^{t-1}_{i,j} + O_{i,j}*(1 - O_{i,j})*\delta^{t-1}_{i+1,k}*Wi_{+1,k,j}$$

$$W^{t}_{i,j,k} = W^{t-1}_{i,j,k} + \eta*\delta_{i,j}*O_{i-1,k} + \alpha*(W^{t-1}_{i,j,k} - W^{t-2}_{i,j,k})$$

$$O_{i,j} = \frac{1}{1+e^{\beta(-(I_{i,j}+bias_{i,j}))}} \tag{4}$$

where

i,j,k	– index of: layer, neuron and weight
t	– point of time
α,β,η	– momentum, beta and learning rate
I,O,A	– input, output and expected value
δ	– error
W	– weight

Paying attention to the given problem the adjustment of the algorithm and coefficients was conducted. The results of the adjustment of coefficients α and η is shown by Fig.7. The most proper choice turned out to be α = 0.6 and η = 0.2.

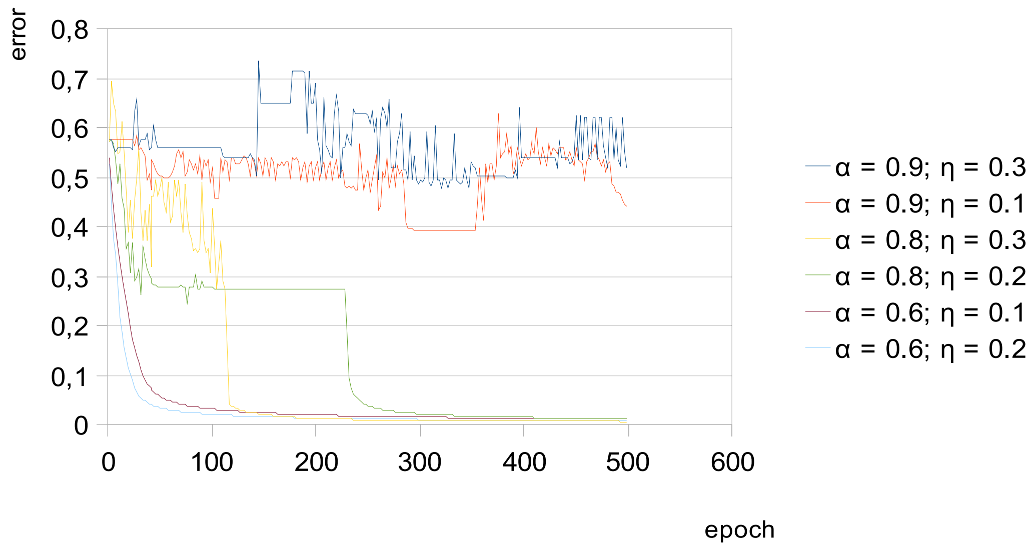

Fig.7. The selection of α and η. Source: Own elaboration

After choosing α and η it is possible to estimate the coefficient β. Fig.8 shows the results of experiments for various β coefficients.

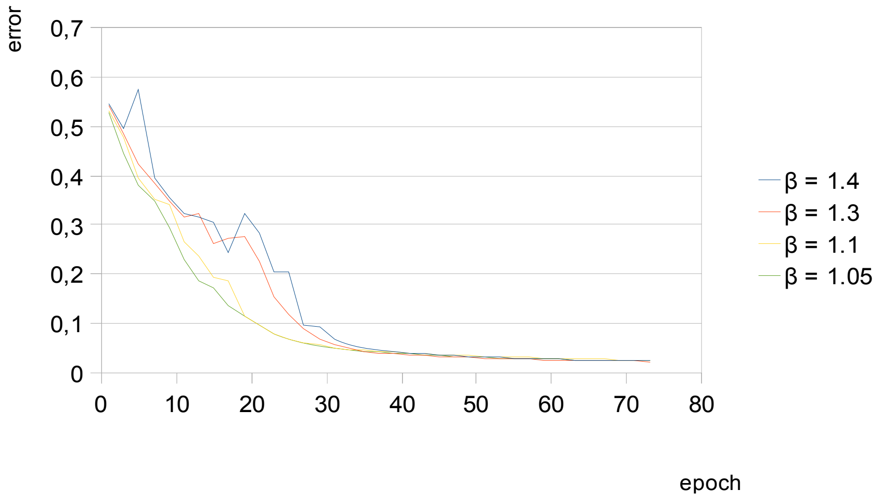

Fig.8. The selection of β. Source: Own elaboration

Analysing the mistake dependence upon the number of epochs it can be noticed from Fig.8 that the most satisfactory results occur for β which is close to 1.

The elaborated program is equipped with the mechanism of measuring out the time needed for learning and recognizing. According to the assumptions from the beginning of this project the computers intended for the general usage were used with Windows 7 Home Premium 64 bit, Dual Core AMD Athlon 64X2 L310 processor and 1.2 GHz RAM 2 GB memory. Satisfactory results of learning were achieved after less than 500 epochs and the time of learning was 5 seconds. The time of recognition fluctuated in single milliseconds. Those two results confirmed the validity of the proposed attitude on the computer equipment intended for the general use which is not very expensive.

The conducted experiments

The given monitoring conception of the system was implemented according to the policy and used for the image analysis coming from the real monitoring. The image referred to the recorded burglaries into the parked cars and thefts (film). Fig.9a) presents the network connected with the real car (first from the left) whereas Fig.9b) shows the network connected with the silver car- on the right. The conception accepted here related to the network with the variable density.

a) b)

Fig.9. Image from the camera with the network put by the monitoring system. Source: Own elaboration

Two examples of the neuron network with one hidden layer were considered. The number of neurons in the hidden layer was established according to pattern (2) while in the adjustment algorithm the accepted coefficients were $\alpha = 0.6$, $\eta = 0.2$ and $\beta = 1.05$. The network was taught on the basis of the patterns worked out according to the human intuition patterning on the trajectories analysed on Fig.2, Fig.3 and Fig.4. The expert's knowledge was used for working out these patterns.

The practical method of adjusting the system

The simulation analysis and the experiments conducted in real situations let to formulate some practical rules of configuring of the elaborated system fir other applications.

- **Determining the network**

  The network can be chosen as homogeneous or with changeable density. Then, the system operator has to indicate the crucial sphere in the image seen by the camera and establish if the network should be connected with the scene being observed or if it should be the object in that scene.

- **Teaching the network**

  In the situation where the current application is very similar to the previous one, the predefined network without teaching can be used. For example, if take into consideration the car park similar to the one presented by Fig.9 etc. If the current situation is totally different from the ones available in the previous study, the proper patterns should be prepared.

  It is important that they can be prepared in two ways:

1) Recording the mock situation or
2) Determine on the basis of the expert knowledge the trajectories by indicating the proper spheres of the network.

Generally, there are even several dozen and that is why it is the task available for the human perception and possible to be released in short time. The system operator has to prepare the trajectories analogous for those presented by Fig.10 b) and c) in the form of the marked spheres of the network.

a)

b)

c)

Fig 10. Course of events in the burglary: a) The network put by the system, b) Suspicious action from the point of view of the in-put neurons, c) The prohibited action. Source: Own elaboration

Conclusions

The presented monitoring system fulfils all the aims established at the beginning of this article which referred to creating low-cost monitoring system which can be combined on the basis of computers intended for general use as well as typical cameras such as network cameras. In addition, the main characteristic feature distinguishing this particular monitoring system from others available on the market was discovered. This feature refers to the ability of take to prevent the prohibited actions and not only to record them. The use of typical neuron network with the standard configuration allowed to receive satisfactory results.

Bibliography

[Pelc, 2008] L.Pelc, B.Kwolek. Activity Recognition Using Probabilistic Timed Automata. In Peng-Yeng Yin: Pattern Recognition. Techniques, Technology and Application. In-Tech, November 2008, pp. 345-360.

[Rutkowski, 2009] L.Rutkowski. Metodyi techniki sztucznej inteligencji. PWN, Warszawa, 2009.

[Tadeusiewicz, 1993] R.Tadeusiewicz. Sieci Neuronowe. Wydawnictwo RM, Warszawa, 1993.

[Wszołek, 2006] W.Wszołek. Metody sztucznej inteligencji w systemach monitoringu hałasu lotniczego. Zeszyt 3/2006 *patenty.bg.agh.edu.pl/graf/Newsletter_AGH_Czerwiec_2011.pdf* [06.2011]

[Zurada, 1996] J.Żurada, M.Barski, W.Jędruch. Sztuczne sieci neuronowe. PWN, Warszawa, 1996

Netography:

http://www.youtube.com/watch?v=pLKjm2uGrU4&feature=related

http://www.ai.c-labtech.net/sn

http://kik.pcz.pl/nn/index.php

http://www.neuron.kylos.pl/pliki/start.html

http://www.willamette.edu/~gorr/classes/cs449/Backprop/backprop.html

Authors' Information

Lucjan Pelc – *lecturer, Rzeszow University of Technology; W. Pola Street 2, 35-959 Rzeszów and The Institute of Technical Engineering, The Bronisław Markiewicz Higher State School of Technology and Economics, Czarnieckiego Street 16, 37-500 Jarosław, Poland; e-mail: lpelc@prz-rzeszow.pl*

Major Fields of Scientific Research: Real'time protocols and systems, Formal description techniques, industrial vision systems

Artur Smaroń – *computer science's student, The Bronisław Markiewicz Higher State School of Technology and Economics, Czarnieckiego Street 16, 37-500 Jarosław, Poland; e-mail: arturo.s@op.pl*

Justyna Stasieńko – *lecturer, The Institute of Technical Engineering, The Bronisław Markiewicz Higher State School of Technology and Economics, Czarnieckiego Street 16, 37-500 Jarosław, Poland; e-mail:* justyna.stasienko@pwste.edu.pl

Major Fields of Scientific Research: Management Information Systems, Business information technology