Research on the development of recurring neural networks

Mikhail Mikheev, Julia Gusynina and Tatyana Shornikova

Penza State Technological University, 1a/11, Baidukova pas./Gagarina str., Penza, 440039, Russian Federation
shornikovat@mail.ru

Abstract. The article studies the development of recurrent neural networks. Various existing approaches to the construction of the architecture of recurrent neural networks, as well as their training, are considered. A new impulse neural model has been built using the so-called reservoir transformations. The research of this model, ranking and classification of its most significant indicators has been carried out. The main problem of pattern recognition is determined. The main indicators of the "black box" of the neural network have been identified. The degree of influence of these indicators on the development of a recurrent neural network has been practically proven. An algorithm for connecting neuron impulses into a single whole has been developed. An example of the effective use of this algorithm for determining the problem of recognizing constantly changing network images is shown.

Keywords: Recurring neural network, pattern recognition, neuron impulse, input and output data, training function, resting potential, dumping potential, synapse.

1 Introduction

The problem of pattern recognition is currently very relevant. There are many ways to solve this problem [1-2]. A very promising way to solve these problems is neural networks, namely, recurred neural networks with feedback. Due to feedback, input parameters instantly spread in the environment of neural networks. This leads to almost 100% learning capacity of networks and to the huge computational potential of recurring neural networks [3-4].

The search for solutions to problems associated with recurring neural networks began to be engaged in the second half of the last century. Now it is known about several dozen recurring neural networks. In addition, there is also a sufficient number of means to address such challenges [5]. Most often, all tasks associated with recurred neural networks are divided into two types - controlled and unmanaged.

* Copyright c 2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-2922/paper005.pdf

Controlled neural networks interact with each other for a while, and then come to the same equilibrium state [6]. This state is a multidimensional point of space that becomes a prototype of these networks. This point is the so-called "memory point" of the process. The nature of equilibrium points can be different: from more specific to more messy. Recurring neural networks are then called oscillatory and disorderly. Such networks lead to one common view of their neurons, which can be applied in problems of hierarchy construction and order of objects [7-8].

Recurrence neural networks of an unmanaged type develop as you like. They are based on the following: what comes at the input must be correctly converted at the output [9]. This principle combines unmanaged neural networks with conventional neural networks. As a result of training, this network distributes weights so that their total amount tends to zero[10].

There are a huge number of approaches to learning neural networks [11]. The bulk of them are approaches related to optimization methods, the main of which is the gradient method. A little stand alone is the method of training the machine. In addition, there are more modified methods of training the network [12].

Consider the algorithm for constructing the architecture of recurring neural networks. We describe the structure of the model, its constituent subsystems and elements, the connections between them. Let us highlight the main parameters affecting the functioning of the model as a whole, as well as the totality of parts.

2 Materials and methods

The neural network contains at its core the main set and a set of readers. Signals are input that, using special transformations, reach a state that helps them get out of the set in the form of functional dependencies. In this case, no pre-training of the network is required. The states into which the recurring neural network falls are displayed by special devices. These devices read the images of the network, comparing them with similar ones, as a result of which they correctly recognize objects. The solution to the problem of image recognition by a recurring neural network is shown in Fig. 1.

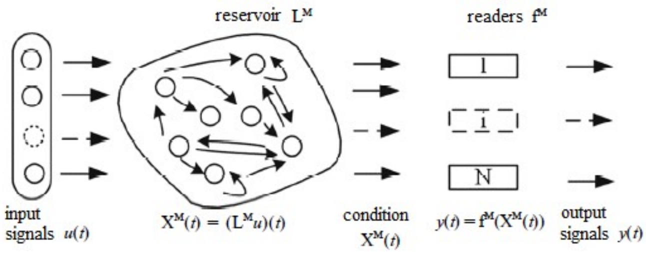

Fig. 1. Black box diagram of a recurring neural network.

The recurring neural network has a pulsed character. Its structure is three-dimensional and obeys the stochastic law. The connection between the network components is as follows [3]:

$$P(a,b) = \min[1, C(a,b) \cdot \exp\left(-\left(D(a,b)/\lambda\right)^2\right)] \quad (1)$$

The formula describes the relationship between two neurons a and b; the distance between them is indicated through $D(a,b)$; the relationship between neurons is expressed through the parameter λ; compactness of relationships represents quantity $C(a,b)$, it depends on the types of neuron, which are most often two.

We investigate in more detail each of their subsystems of recurring neural networks.

The first kind of neurons is associated with impulse characteristics of the process. Messages are sent to the network input both in analog form and in pulse form. In addition, the network is in an excited state due to external noise interference of the system [12, 13].

The model of this kind itself is very easily characterized by differential equations of the first or second orders. The most common dependency with which the study model is depicted is as follows:

$$C\frac{dv}{dt} = \frac{v}{R} + \frac{v_{rest}}{R} + i_{ext} \quad (2)$$

Here v_{rest} characterizes the resource neuron; R - this is the physical resistance of the device; i_{ext} - this is the pulse applied to the network input.

The process of excitation of the network is as follows: after the resource of the neuron has approached its critical value θ, powerful pulse emission occurs. Neuron resource decreases to magnitude v_{rest}, the neuron becomes immune to external interactions, and the state of the neuron itself is described by the magnitude T_{refrac}.

These models are a symbiosis of Izhikevich models [5] and models of the Hodgkin-Huxley type. In addition, they can be a modification of other types of models. Typically, an analytical model entry has the differential equation itself or a system of differential equations and a boundary condition. The differential equation has the form:

$$\begin{cases} \dfrac{dv}{dt} = 0,04v^2 + 5v + 140 - u + i_{ext}, \\ \dfrac{du}{dt} = a(bv - u). \end{cases} \quad (3)$$

And the boundary condition looks like this:

$$\text{if} \quad v \geq 30, \quad \text{then } v = c, u = u + d \quad (4)$$

Here v - is the neuron resource; magnitude u describes oscillatory processes; i_{ext} - this is the input value; t - this is a time characteristic; a, b, c, d - different components of the model.

The configuration of recurring neural networks consisting of pulsed neurons is quite different. It depends on the number of neurons, as well as the timing of the pulses. For example, more complex models such as Hodgkin-Huxley consist of an average of 50 neurons, Izhikevich models already have about a thousand neurons, and excitatory models contain many more neurons [6].

Recurved neural networks described by pulses have synthetic connections. These links have a number of drawbacks.

First, only the impulse drives the response of neurons, the rest of the time they are passive. Second, impulse response always occurs with little delay in time.

Current in these models i_{post}, following pulses, can be described by an indicative function, which decreases monotonously. This function has the form:

$$i_{post}(t) = w_{chem} \sum_{i} 1(t - t_i - \tau_{delay}) \exp\left(\frac{t - t_i - \tau_{delay}}{\tau_{syn}}\right) \quad (5)$$

$$\frac{dv}{dt} = f(v, par) + k i_{post}(t) \quad (6)$$

Here w_{chem} characterizes the relationship of pulses; t_i - this is the time characteristic of the i - pulse; τ_{delay} - this is a late function; $1(t - t_i - \tau_{delay})$ – a multilevel Heaviside dependency that changes its value from zero to one at each time point $t_i + \tau_{delay}$; time parameter τ_{syn} describes the gradient of decrease of the indicative function.

Sometimes more complex models are used, taking into account other types of pulses. So, the previous moments of time take into account the so-called developing pulses [7], which process the input and output displays taking into account their prehistory. So-called flexible pulses [6] are able to adapt their flexibility according to the Hebb algorithm, which allows increasing the amplitude of pulse propagation where neurons show their hyperactivity.

To recode information from the inputs and outputs of a recurring neural network, there are many ways. The main ones process information from continuous to discrete or from discrete to continuous using special converters. These transducers are based on synchronous methods. These devices analyze the development of the network and significantly simplify the process of solving problems.

3 Results

To solve the problem, let's classify the parameters of the neural network.

Mostly parameters are divided into three classes. The first class characterizes the input characteristics of the network. They consist of images that are compared to known or similar presented. These features are distinguished by their properties. Depending on their properties, features give greater or lesser convergence to zero during pattern recognition.

The second class of parameters is associated with emitted pulses of the neural network. They also have certain properties that allow you to recognize images to one degree or another.

You can construct a mathematical function of the effect of pulses on pattern recognition parameters. It will depend on the number of network elements $n_{nrn} = n_x n_y n_z$, neuronal interconnectedness parameter λ, number of connections from network input to network output p_{vh}, neuronal concentrations $C(a,b)$, cognitive weights $W_{\text{int}}(a,b)$ and non-cognitive $W_{vh}(a)$ parameters, distribution of number of different-level neurons p_{inh} and noise immunity p_{noise}.

Models can be both simpler and more complex. Noise resistance i_{noise} and boundary conditions v_0 are considered the main characteristics of both models.

IaF models are divided into two types: with and without loss. The main characteristics of such models are as follows: time constant τ, moment of unresponsiveness T_{refrac}, static v_{rest}, ejection v_{reset} pulses, critical value θ. Izhikevich models are defined by characteristics a, b, c, d. More complex models define their parameters more strictly.

The main characteristics of their interaction are described by the following values: target function $W_{\text{int}}(a,b)$ and delay parameter τ_{delay}. Electrical pulses occur instantaneously, so $\tau_{delay} = 0$. Dielectric pulses appear wave-like and are described by a time constant τ_{syn}. Impulses with development are characterized by a tuple U, D, F. Flexible pulses describe their objective function $w_{\max}$ as the smallest difference of the synapse $\Delta w_{\min}$ and using the time characteristic ΔT_{forget}.

The specificity of flexible pulses is that learning the neural network in this case can occur without a teacher. Then the following values must be entered into the recurring neural network: time characteristic T_{learn}, learning dependency $f_w(\Delta t)$ and its components. The constraint is best approximated by an indicative function that contains values X_{0LTP}, T_{LTP}, X_{0LTD}, T_{LTD}, a hyperbolic function containing values A_{LTP}, A_{LTD}, $\Delta T_{\max}$, $\Delta T_{\min}$ or a Gauss curve consisting of values C, β, $F_{\max}$, B, $\Delta T_{\max}$, $\Delta T_{\min}$.

The third class of parameters is associated with special devices, the so-called network readers. These devices can perform various functions and tasks. So, some of them process the input signal into a stream, which allows you to further solve the problem of recognizing an object. Others change the nature of the signal from discrete to continuous. Others monitor output flow conversions. Fourth calculate the necessary values of the flow of neurons. Ultimately, all of them are aimed at solving the problem.

All the above types of pulses of the recurring neural network can be arranged according to the contribution that they leave when solving the problem of pattern recognition. The input data contains the signal supply method, the types of transformations, the nature of the transformations, the complexity of the network, the dimension of the signal display.

Method of signal supply depends on initial conditions and includes analog and pulse signals. Signal conversion takes place according to a kind of algorithm and contains various conversion parameters. The nature of the transformations can be fast and slow, random and deterministic. Network complexity may or may not be present at all, or it may be simple and complex enough. The dimension of the signal display is classified into a dimension in space depending on the inputs and a dimension in time depending on the duration of the signal presentation.

The "black box" of the neural network itself is characterized by the connectivity of neurons, the type of components, the type of connections, as well as the function of learning.

Connectivity is divided into internal, depending on the midline, communication density, communication strength, percentage of suppressive elements, noise in the network, and external, depending on the percentage of connections of incoming neurons and the strength of connections of neurons. Network element types depend on neuron model. The types of connections are due to the synapse model and depend on the delay and strength of the connection. In addition, they are of electric form (when the delay is zero) and static form, characterized by a constant attenuation time. The training function depends on the training period and is ranked by exponential, threshold, and Gauss curve.

Neuron models are models of integration (or excitation), Izhikevich models, or other more complex species. The synapse model is divided into chemical-dynamic (used at rest) and chemical with plasticity.

Neural network reader performs functions of pulse analysis, classification and clustering.

There are completely different neurons on different layers of the network. To put them together, you need to put together over thirty parameters. The most optimal parameters should be given as many feature values as possible, but the most invaluable and weak ones should be given at most one feature value.

The ranking of the parameters of the recurrent neural network made it possible to determine the approach to their analysis. There are interdependent parameters, and there are isolated from each other. Interdependent parameters need to be correlated together, but isolated parameters can be correlated separately. In addition, some parameters can have a stronger effect on the learning process of the network, while others less. Let's designate them as variable (the most important parameters), adjustable (less important parameters) and fixed, which makes no sense to vary.

The parameters of the first group are characterized by internal connectivity $\lambda \leftarrow C \leftarrow W_{\text{int}}, P_{inh}$, external connectivity parameters P_{vh}, W_{vh}, noise parameters p_{noise}, i_{noise} and a reader $t_{readout}$.

The second group is determined by the number of neurons n_{nrn}, IaF parameters of neurons τ, T_{refrec}, parameters of Izhikevich neurons a, b, c, d, synapses τ_{syn}, training time T_{learn}.

The third group has the following features: the number of inputs n_{vh}, the delay parameter in synapses τ_{delay}, the parameters of dynamic synapses U, D, F, the parameters of synapses with plasticity $\Delta w_{\min}$, $\Delta w_{\max}$, ΔT_{foget}.

The inner part of the "black box" of the neural network is set by such parameters that could recognize a moving image. This task is more complex and is part of a common object recognition task. The main idea of the problem is as follows: the network must compare the images with those already known to ensure that the recognition error tends to zero. This error is generally calculated as the deviation between the output images and the reference ones [12].

In the model under consideration, a complicated recurrent neural network takes place - this is due to the peculiarities of the image at the input of the network, its structure.

Thus, it is necessary to introduce the tuning criteria for the neural network being built. To take into account the dynamics of the neural network, the parameters must be of long duration, but with the final result. And to take into account the speed of the neural network's reaction to pattern recognition, it is necessary to consider the parameters reflecting the difference in the neural network's response to images supplied to the input.

4 Discussion

For experimental testing of the constructed recurrent neural network, a certain procedure was formed. At the first stage, a test set of input images that are uncorrelated with each other is built. Further, the possible boundaries of the reservoir parameters were indicated, experimental studies with their combinations were carried out, and thus quality indicators were obtained. At the same time, the parameters were divided into fixed and dynamic, depending on their degree of influence on the indicator. According to this algorithm, all quality indicators were checked until all the relationships between the parameters were obtained.

During the experiment, two studies of a recurrent neural network were carried out: with integration and actuation neurons and Izhikevich neurons [14].

The results of the first study confirmed the assumptions about taking into account the dynamics of the neural network. It passes from fading memory to non-fading memory, and a decrease in the noise boundary leads to a decrease in the characteristic. It was found that if there is no suppression, then persistent memory appears at a connectivity from 0.02 to 0.025.

In the second study (Izhikevich neurons), no noise was applied to the input of the neural network, this is due to the characteristics of the neurons for which the input signal is important. It was fed with a delay of 0.1 s from the starting point. Another important point is the study of the trigger pulse of the neural network, which propagates to each neuron. And here the conditions of the reference point are important, they correspond to the equilibrium points.

If we consider the situation of the absence of suppressive neurons, then some patterns emerged here:

1) The speed of reaction of the neural network depends on the speed of the sequential increase in the supplied pulse frequency and their attenuation in the end;

2) The typed reaction speed of the neural network is uneven, since the impulse created by the neurons at the input decays for a few milliseconds, and then the network comes into dynamics;

3) The completion of the excitation of impulses by neurons is influenced by the scheme of a recurrent neural network, its division into regions. Since, usually, the fading of the network reaction rate occurs in smaller connected areas that unite interconnected neighboring neurons [15].

Consider the strength of the influence of suppressive neurons. A recurrent neural network responded only to the input impulse if it had suppressive neurons in the circuit and did not consist of excitatory neurons. Consequently, the activity of the neurons themselves was close to zero. In the presence of excitation and suppression neurons with different quantitative composition in the neural network scheme, the neural network response rate was explained by their percentage, the strength of their connection W and the density value C [16-17].

Thus, the neural network responded to an increase with an increase in the proportion of excitation neurons and, accordingly, a decrease in the proportion of suppression neurons. At the same time, the following features of the speed of the neural network's reaction to the initial impulse were observed:

1) The reaction of suppressive neurons approached completion with an increase in the strength of the connection W between suppressive neurons;

2) The reaction rate of excitation neurons directly influenced the response of suppression neurons in the case of transmission of communication from excitation neurons to suppression neurons;

The nature of this process is explained by the fact that the input impulse sets two types of impulses: one goes from excitation neurons and to subsequent neurons, the other - from suppression neurons and further. In this case, there is no connection between excitation neurons, therefore the impulse does not propagate, but there is a connection with suppression neurons, therefore, the impulse can go to them. Accordingly, if the connection between these types of neurons is interrupted, then the reaction rate will be of a different nature.

By themselves, suppression neurons are inactive, but when there is a connection with excitation neurons, they begin to send impulses and come into excitement. At the same time, excitation neurons do not affect themselves like these. This suggests that the slow response rate of excitation neurons generates an enhanced response of suppressive neurons.

3) The reaction rate of the neural network is most strongly manifested in the connections between the suppression and excitation neurons, as well as in the inverse relationship. In this case, a slowdown of suppression neurons occurs if there is no relationship between excitation and suppression, and the suppression neurons themselves do not affect the excitation neurons in any way. In the case of backward propagation of communication, to excitation from suppression, the activity of excitation neurons remains at the same level, that is, suppression neurons do not affect them. If we judge the magnitude of the network reaction speed by the action of these types of communication directions, then in general it decreases [18].

The process of studying the further dynamics of the network showed that the neural network built on these principles reacts at the same speed to irritating impulses, this was especially clearly manifested after 0.4 s from the start of the impulse.

Research has shown that:

- Neurons of suppression and excitation have a joint effect on the reaction rate of the neural network after the impulse;
- The state of rest and coming into the activity of the network directly depended on the speed and strength of the stimulation of this type of neural network, the time of excitation is in direct relationship with the recovery time of the network [19];
- The number of acting impulses does not have a significant effect on the neural network, since the launch of one impulse has already brought the recurrent neural network into activity for a fraction of a second;
- The reaction of the neural network to the next impulse is explained by the time of the given triggering impulse. That is, the network reaction speed will be practically absent if the excitation pulse is applied after the network is restored or during its initial period. In the case of an impulse, the closest to the end of the recovery period of the network, the reaction rate will be opposite to the previous state;
- The network reaction speed directly depends on the density of connections. In the presence of areas with an increased density of connections from excitation neurons to similar ones, the reaction of the entire network will be similar in character to the reaction between these neurons. That is, excitation neurons determine the nature of the network activity, and only they determine its further dynamics [20-21];
- The relationship between suppression and excitation explains the constancy of the active state of the neural network. If the proportion is higher in the direction of excitation links, then the reaction will manifest itself in a solitary and long-term state.

Next, let's look at the effect of noise. We will feed noise of varying intensity to the input of the neural network. And it can be noted that with a value of the link strength indicator of 0.01 and a uniformly distributed ratio of links, the dynamics of the neural network will be characterized by a random distribution of parameters.

In this case, several points can be highlighted:

- The more pulses we launch into the network, the greater the response to noise we will get;
- As for any random process for the network, there are limiting values of the noise power indicator, crossing the boundaries of which the network ceases to perform the recognition function;
- A stable value of noise at a low level allows to increase the reaction speed of an impulse recurrent neural network. This dependence is explained by the fact that noise forces the network to be in a constant state of readiness to recognize input patterns.

A series of experiments made it possible to create a dynamic pattern recognition scheme:

1) Setting the method for submitting data to the input;

2) Determination of dynamic and static parameters of the network, types of neurons and synapses, selection of dynamic parameters to achieve the optimal research goal;

3) Selection of parameters and conversion schemes based on the dynamics of the network;

4) Evaluation of the obtained indicators and their comparison with the parameters specified in the research objectives.

5 Conclusion

For the investigated neural network with feedback between neurons, different layers of neurons determined the leading characteristics, their order in the general scheme. Parameters have been divided into several types: dynamic, static, and custom. This division is explained by the degree of the contribution made to the operation of the neural network, the nature of their interaction, the network's response to communications and work with readers. At the same time, for each characteristic, a boundary value is indicated that explains the speed of the network's reaction.

The preparation procedure for solving the problems of pattern recognition is indicated, the procedure for selecting the data supplied at the input and received at the output of the neural network is determined. This is especially true for manually untagged images.

It is proposed to distinguish more complex and less complex recurrent neural networks by the type of indicators that form the network. Based on the complexity of the practical problem, it is necessary to model the corresponding degree of complexity of the neural network used for the solution.

The practical part of the study was aimed at comparing the network parameters that make up the model of impulse neural networks with integration and trigger neurons and that are dependent on the input signal by neurons. As a result, criteria were obtained that explain the reaction rate of the neural network according to the values of the indicators that make up the network diagram, the strength of the connection, the ratio of suppressive and excitatory neurons, the presence and value of the noise indicator.

The above analysis can be used to solve a number of computer vision problems even before the construction and training of a suitable network, at the stage of its modeling. This approach will significantly reduce time and financial costs.

References

1. Golovko, V.A.: Neural Networks: Training, Organization and Application: Textbook. Manual for Universities. IPRZhR, Moscow (2000).
2. Batyrkanov, J.I., Kudakeeva, G.M., Subankulova, J.Zh.: Recognition of visual images: an approach based on reference images and training, Ogaryov-Online 15 (2017).
3. Abdugulova, Zh.K., Kudaibergenov, A.K., Aizhanov, A.K.: Problems and prospects of the development of pattern recognition. Actual research in the modern world, 6–1(38), 5–10 (2018).
4. Oligova, M.M.: The use of neural networks in pattern recognition problems. Academy of Pedagogical Ideas Novation. Series: Student Scientific Herald, 1, 91–94 (2019).
5. Bezhin, N.V.: Neural networks in the problem of pattern recognition. Scientific community of students. Interdisciplinary research, 72–77 (2019).
6. Bazhenov, E.N.: Analysis of the use of neural networks and geometric-structural approach for pattern recognition. Modern trends in the development of science and technology, 3–4, 10–12 (2017).
7. Fedoseev, A.A., Fryshkina, E.A.: Pattern recognition. Scientific notes TOGU, 9 (2), 475–479 (2018).
8. Zemlevsky, A.D.: Study of the architecture of convolutional neural networks for the task of pattern recognition. Bulletin of science and education, 6(30), 36–43 (2017).
9. Mikheev, M.Yu., Gusynina, Yu.S., Shornikova, T.A.: Building neural network for pattern recognition. In: 2020 International Russian Automation Conference (RusAutoCon), 357-361. Institute of Electrical and Electronics Engineers, NY (2020).
10. Medvedik, A.D., Volkov, N.V, Konyukhovskii, S.M.: Evaluation of information content of moment invariants used in pattern recognition. Modern information and electronic technologies, 1 (17), 85-86 (2016).
11. Enweiji, M.Z., Lehinevych, T., Glybovets, A.: Cross-language text classification with convolutional neural networks from scratch. Eureka: Physics and Engineering, 2, 24-33 (2017).
12. Drokin, I.S.: About an algorithm for consistent weights initialization of deep neural networks and neural networks ensemble learning. Vestnik of Saint Petersburg University. Applied mathematics. Computer science. Control processes, 4, 66-74 (2016).
13. Mikheev, M.Yu., Gusynina, Yu.S., Shornikova T.A.: Problems of using neural networks. J. of Phys.: Conf. Ser., 1661, 012104 (2020).
14. Limontsev, D.S.: Texture recognition algorithm using deep neural networks. 2019 Information Technologies, Energy and Economy Conference, 303 – 306 (2019).

15. Mikheev, M.Yu., Gusynina, Yu.S., Shornikova, T.A.: Construction of intellectual informative systems. In: CEUR Workshop Proceedings, 2843. CEUR-WS Team, Aachen (2021).
16. Goldstein, M.A.: Use of convolutional neural networks in the problems of stylization and synthesis of textures. Alley of science, 2 (9), 849-855 (2017).
17. Baranov, K.A., Chaychits, N.N.: Recognition of images and data processing using neural networks. Actual directions of scientific researches of the xxi century: theory and practice, 6 (42), 37-38 (2018).
18. Siyakina, V.V., Salakhutdinov, E.R., Shubin, A.V, Erokhin, A.A.: Application of convolutional neural networks for pattern recognition. Modern Science, 6 (1), 235-238 (2019).
19. Verzun, N.A., Kolbanev, M.O., Omel'yan, A.V.: Perspective technologies of infocommunication interaction. LETI, St. Petersburg (2017).
20. Sovetov, B.Ya., Cekhanovskij, V.V.: Information technology. Urait, Moscow (2016).
21. Mikheev, M.Yu., Gusynina, Yu.S., Shornikova, T.A.: Recognition of textures using instant features and neural network methods. Bull. of scientific centre of children safety, 4 (46), 137-146 (2020).