Application of deep learning methods to ensure the information security of the MSAB digital system in the railway transport of the Republic of Kazakhstan

Malika Sagitzhanova[1,*,†], Kanibek Sansyzbay[2,†], Yelena Bakhtiyarova[2,†] and Teodor Iliev[3,†]

[1] *Mukhametzhan Tynyshbayev ALT university, Almaty, 050000, Kazakhstan*
[2] *International Information Technology University, Mans St. 34/1, Almaty, 050040, Kazakhstan*
[3] *University of Ruse, Ruse, Bulgaria*

Abstract

Digitalization of railway traffic control systems requires the use of modern approaches to information security (IS). One of the promising technologies is the use of Deep Learning methods, which automatically detect anomalies and potential attacks in complex automated control systems. Attention is given to choosing the most suitable models for detecting attacks in the KZ-MSAB-MA digital semi-automatic blocking system under development.

Keywords

information security, deep learning, neural network models, national digital MSAB system

1. Introduction

The railway infrastructure is characterized by a high degree of distribution, the presence of long stretches, as well as a significant number of autonomous technical devices located on open sections of track. These features create unique challenges for information security, especially in the context of the transition from isolated relay systems to digital architectures with the possibility of remote control and monitoring [1].

Key elements susceptible to cyber threats include alarm, centralization, and lockdown systems that control the safe movement of trains. Despite the fact that such systems have traditionally operated in closed networks, there is a tendency to integrate them into corporate IP infrastructures. This expands the capabilities of dispatch control, but also increases the attack surface [2].

An additional risk is created by the use of wireless communication channels (TETRA, GSM-R, LTE-R, Wi-Fi), which are used for communication between the train and the infrastructure, as well as between automation devices. Such channels are vulnerable to interception, data substitution, and man-in-the-middle (MITM) attacks, which makes the task of detecting attacks especially important [3].

In general, threats can be divided into several key categories:

- attacks on data integrity — changing control commands or state parameters (for example, false signals about the release of a path);
- Attacks on accessibility — blocking of communication channels (DoS/DDoS), disabling control nodes;
- Privacy attacks — interception of control commands or route data;

[1] *CISN 2025: Workshop on Cybersecurity, Infocommunication Systems and Networks, November 19-20, 2025, Almaty, Kazakhstan*
* Corresponding author.
† These authors contributed equally.
mzhandosovna77@gmail.com (M. Sagitzhanova); k.sansyzbai@iitu.edu.kz (K. Sansyzbay); y.bakhtiyarova@iitu.edu.kz (Y. Bakhtiyarova); tiliev@uni-ruse.bg (T. Iliev)
0000-0002-3333-5830 (K. Sansyzbay); 0000-0001-8735-7683 (Y. Bakhtiyarova); 0000-0003-2214-8092 (T. Iliev)

- Malicious software injection — gaining access to software platforms (for example, SCADA), installing hidden programs;
- Interference with wireless interfaces — signal substitution or distortion, command interception.

Gradual, subtle attacks that simulate the "normal" behavior of the equipment pose a separate danger. They cannot be detected by simple filtering and threshold control methods.

Under these conditions, the use of intelligent analysis technologies, in particular deep learning methods, can provide preliminary protection and timely detection of anomalies in the operation of key components of the railway IT infrastructure. These systems are capable of processing large amounts of data about the operation of equipment and network traffic, identifying unusual or suspicious patterns that are not detected by traditional security methods. This is especially important for automated railway systems, where rapid response and accurate threat detection are critical [4].

2. The architecture of the MSAB digital system with the integration of the DL module

The national digital MSAB system being developed is designed to automate traffic control on straight sections of the crossing. Its use makes it possible to increase the speed of trains to 200 km/h with good track condition. The system is based on two half-sets, each of which includes a KZ-MSAB-MA controller and expansion modules (DI/DO) located at the stations [5].

The controllers interact with external control panel devices (rail circuits, axle counting devices, switches, etc.) and control the alarm based on data received through digital inputs. Control commands are transmitted through digital outputs to the relay. A data exchange channel using the SIL4 security level protocol is used for communication between the controllers. The channel can be built on the basis of copper lines, Fiber-Optic Communication Lines or a radio channel, depending on the infrastructure conditions. Figure 1 presents the schematic diagram of the integrated system KZ-MSAB-MA under development, featuring intelligent control functions within a distributed architecture [6].

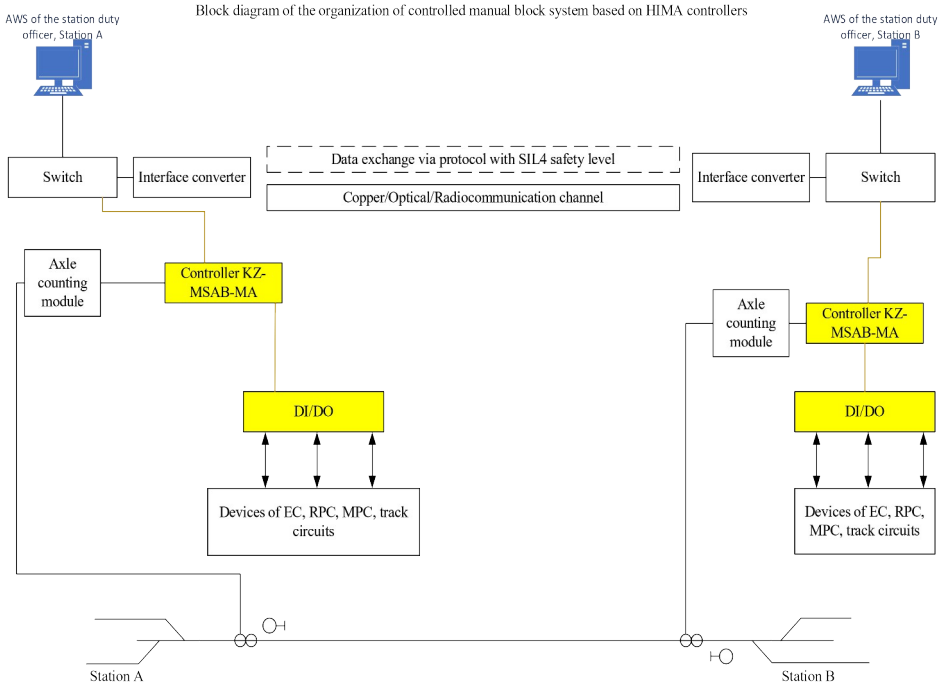

Figure 1: Approximate architecture of the developed prototype of the integrated distributed modular digital semi-automatic blocking system KZ-MSAB-MA [6].

3. Results

The description of the algorithm of the MPAB system for railway transportation (from the departure signal to the entrance signal) is presented in the article using the programming languages FBD (Function Block Diagram) and ST (Structured Text).

The exterior of the MPAB modular architecture cabinet is shown in Figure 2.

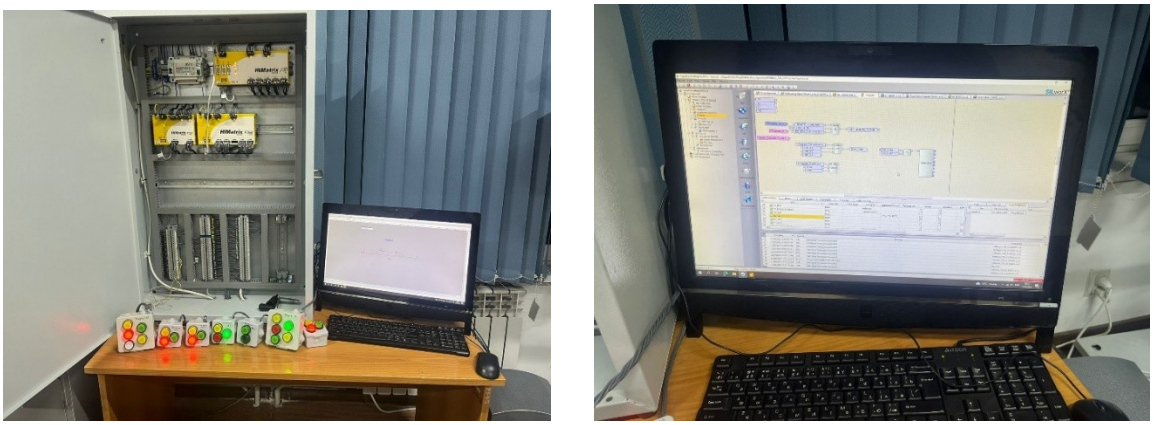

Figure 2: View of the MPAB modular architecture cabinet.

The SILworX environment (HIMA, Germany) was chosen to develop an algorithmic framework for modular hardware and software locking (MPAB). The choice of this environment is determined by its compliance with the Safety Integrity Level (SIL) T3 according to the international standard IEC 61508-3. This, in turn, ensures compliance with the strict safety requirements established by the European standards CENELEC and the Russian technical regulation TR/TC 003/2011 "On the safety of railway transport infrastructure".

The highly reliable HiMatrix F3 DIO 20/8 and HiMatrix F3 DIO 8/8 relay blocks belonging to the 1st reliability class are used as actuators.

Safety conditions and input/output commands for managing key railway infrastructure facilities (routes, traffic lights, arrows) are conveniently formalized as one-dimensional arrays [20]. In the context of security, the logical '0' is accepted as an indicator of the safe state of the system.

The software architecture is based on an object-oriented approach: each physical object of the station (traffic light, arrow, or track section, as shown in Figure 3) is implemented as a separate Functional Block or function. Each of these blocks corresponds to a graphic designation adopted in the development of functional circuits of electronic devices.

Figure 3 shows the algorithm for setting the route to the first path at the signal of traffic light H. The process consists of two main sequential actions:

1. Verification of the route installation conditions: assessment of the possibility of forming a route without directly switching the permitting traffic light reading;
2. Route locking: fixing the selected route in the centralization system.

Traffic lights are among the key floor-mounted devices of railway automation that ensure safe and regulated train movement.

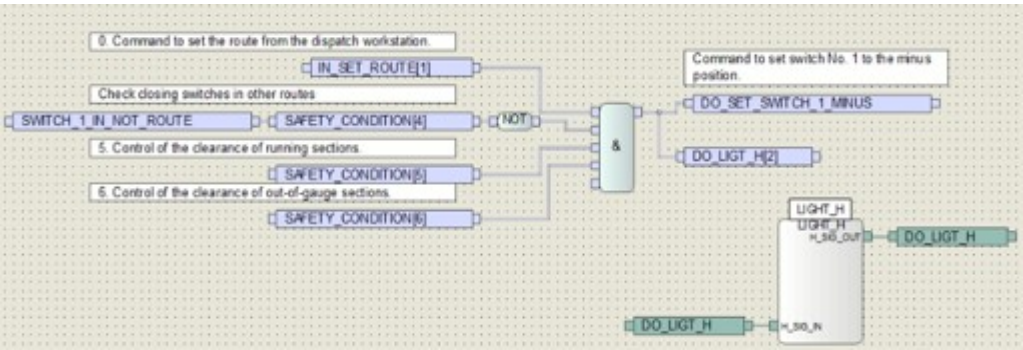

Figure 3: Algorithm for setting a route to the first track based on signal H.

As an illustration of the software implementation of the control, the algorithm for turning on the traffic light H, implemented in the Structured Text (ST) programming language, is presented. This program code, labeled as (Function to set input signal H), is integrated and used inside the LIGHT_H function block.:

```
IF AND (H_SIG_IN[1],H_SIG_IN[4], NOT(H_SIG_IN[2]),NOT(H_SIG_IN[3]),NOT(H_SIG_IN[5]))
THEN (*Processing set two yellow signal command*)
H_SIG_OUT[1] := TRUE; (*YellowUp*)
H_SIG_OUT[2] := FALSE; (*Green*)
H_SIG_OUT[3] := FALSE; (*Red*)
H_SIG_OUT[4] := TRUE; (*YellowDown*)
H_SIG_OUT[5] := FALSE; (*White*)
ELSIF AND (H_SIG_IN[2],
NOT(H_SIG_IN[1]),NOT(H_SIG_IN[3]),NOT(H_SIG_IN[4]),NOT(H_SIG_IN[5])) THEN
(*Processing set green signal command*)
H_SIG_OUT[1] := FALSE; (*YellowUp*)
H_SIG_OUT[2] := TRUE; (*Green*)
H_SIG_OUT[3] := FALSE; (*Red*)
H_SIG_OUT[4] := FALSE; (*YellowDown*)
H_SIG_OUT[5] := FALSE; (*White*)
ELSE
H_SIG_OUT[1] := FALSE; (*YellowUp*)
H_SIG_OUT[2] := FALSE; (*Green*)
H_SIG_OUT[3] := TRUE; (*Red*)
H_SIG_OUT[4] := FALSE; (*YellowDown*)
H_SIG_OUT[5] := FALSE; (*White*)
END_IF;
```

To monitor the status and detect attacks, it is proposed to integrate the DL module into the system architecture. DL is implemented as a passive intrusion detection system (IDS), which is located at the server level of the automated workstation of the duty officer or as an intermediate element on a network node between stations. The security module receives copies of network traffic between controllers and analyzes the following parameters: time characteristics of packets, DI/DO command sequences, the status of axis counting devices and responses from rail circuits [7]. Figure 4 shows a system with a built-in DL-based information security module.

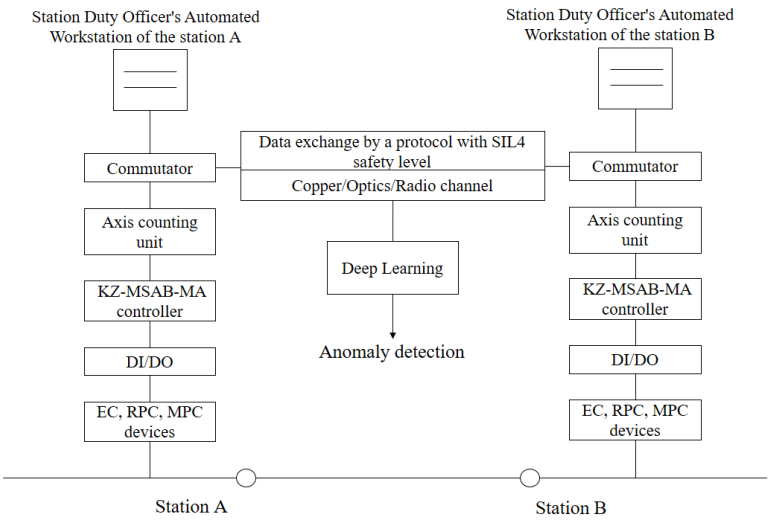

Figure 4: Structural diagram of the DL module integration into the architecture of the system under development.

4. An overview of deep learning techniques for attack detection

4.1. Autoencoders

Automatic encoders are effectively used in information security systems of the railway industry, especially in anomaly detection and attack tasks. The main idea is the ability of the AE to learn from normal data, identifying deviations from typical behavior.

One of the key uses is to detect anomalies in telemetry and network data coming from trains, signalling devices, routing systems, and other infrastructure components. The AE is trained on correct (normal) data, and when anomalies occur — for example, as a result of an attack or technical failure — the model detects a significant discrepancy between the input and output, which allows a quick response to the accident.

In addition, automatic encoders help reduce dimensionality and clear data from noise, which is especially important when analyzing large streams of information in real time. In particular, noise-canceling systems increase the resistance of systems to distortions that occur, for example, due to interference in data transmission channels.

AE is also used in network traffic monitoring systems between trains and control centers. Deviations in the traffic structure may indicate intrusion attempts, such as the introduction of malicious code, interception, or substitution of control commands [8].

4.2. Recurrent neural networks (RNN)

Recurrent neural networks are particularly useful in analyzing sequential data, which makes them effective in monitoring and detecting anomalies in time series typical of railway systems. Such data may include telemetry from rolling stock, traffic control system logs, and safety events.

RNNs are able to take into account the temporal context, which makes it possible to identify hidden patterns and anomalies that do not occur simultaneously, but as a result of gradual changes in the behavior of the system. This is critical, for example, when detecting prolonged attacks or degradation of system components.

Due to the ability to predict future conditions based on history, RNNs are used for early warning of potential incidents, including cyber attacks, instability of communication channels and failures in alarm systems [9].

4.3. Convolutional neural networks (CNN)

Convolutional neural networks are traditionally used for image processing, but they are also successfully used in data analytics and cybersecurity, especially when working with the representation of network or temporal data in the form of matrices or "maps".

In the railway industry, convolutional neural networks (CNNs) are used to analyze network traffic, video streams from surveillance cameras, and event logs transformed into visual representations. They allow you to detect anomalies, unauthorized access, and other threats with high accuracy, including those that are difficult to detect using traditional methods.

4.4. Hybrid models

Hybrid neural network models combine the capabilities of various architectures, such as CNN, RNN, and autoencoders, providing comprehensive threat analysis. CNNs extract spatial features, RNNs process temporal dependencies, and autoencoders detect anomalies.

In the railway sector, they are used to monitor IT infrastructure, predict attacks, and intelligently filter events. Such models are particularly effective at detecting complex threats, including APT attacks, due to their high accuracy and adaptability to diverse data sources [10].

4.5. Self-supervised models

Self-supervised learning models make it possible to identify threats without the need for manual data markup, which is especially important in conditions of limited information about cyber attacks.

They effectively analyze logs, telemetry, and network traffic, identifying hidden patterns of behavior and adapting to new threats. Such models can be continuously updated and can be easily integrated into hybrid solutions, enhancing the protection of cyber-physical systems of the railway infrastructure.

Table 1
Comparative table of application of neural network models

Model	Main application	Advantages	Data type	Examples of tasks
Automatic Encoder (AE)	Anomaly detection, data compression, noise reduction	It does not require marked-up data, it identifies hidden threats	Telemetry, network traffic	Deviation detection, pre-filtering, data recovery
RNN (recurrent networks)	Time sequence analysis, prediction of system behavior	Takes into account the time context, suitable for long dependencies	Time series, logs, signals	Monitoring of successive attacks, predicting failures
CNN (Convolutional networks)	Visual and structured data processing	High accuracy in pattern analysis, suitable for transformed data	Images, videos, feature matrices	Analysis of video streams, network traffic, visualization of logs
Hybrid models (CNN+RNN, etc.)	Comprehensive threat analysis, combining spatial and temporal features	High accuracy and adaptability, resistance to complex threats	Combined data	APT (Advanced Persistent Threat) attacks, multi-layer traffic analysis, forecast + classification

A comparison of the models showed that the best results are achieved using hybrid neural network architectures that combine the capabilities of various approaches (for example, convolutional and recurrent networks). Such models allow comprehensive analysis of both spatial and temporal characteristics of data, providing high accuracy and stability when dealing with diverse threats. Next, we will consider convolutional and recurrent neural networks as key deep learning architectures used to analyze complex data in information security tasks.

5. CNN+RNN model

The integration of artificial intelligence (AI) and deep learning methods, such as convolutional neural networks (CNN) and recurrent neural networks (RNN), has become widespread in almost all areas of modern life. These technologies are used to optimize medical care, diagnose diseases and predict their development, as well as to improve financial analysis and decision support.

AI-based approaches demonstrate high versatility and effectiveness, which has led to their application in such diverse fields as agriculture, finance, and healthcare. The CNN-RNN model aims to increase security by implementing reliable mechanisms for detecting and preventing emerging threats in real time. The CNN-RNN architecture combines the advantages of both networks: CNN efficiently extracts spatial features, whereas RNN specializes in analyzing time dependencies in data sequences. The processing process includes two main stages. At the first stage, time series of input data, designated as (x0, ..., x(t - 2), x(t – 1)), are fed into the CNN convolutional layer to extract spatial features. Then, at the second stage, the obtained features are transmitted to the recurrent RNN network, which predicts the sequence of future values (θx(t) , θx(t + 1), . . . , θx(t + m)) m steps ahead [8].

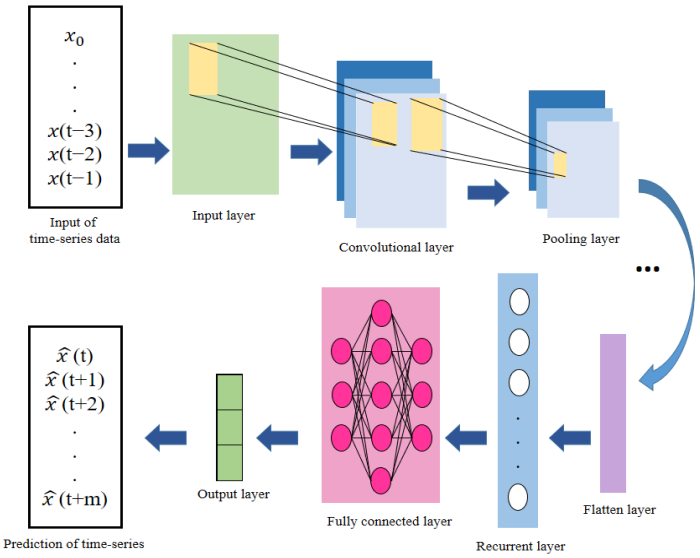

Figure 5: Schematic structure of CNN-RNN.

Thus, the proposed hybrid model is a promising tool for data mining, capable of providing sustainable protection against threats through a combination of spatial and temporal modeling.

In particular, the CNN-RNN model consists of two stages:

Stage 1: The processed data is entered into the CNN structure, which consists of three levels: i) The convolutional layer uses filters to create a feature matrix; ii) The combining layer, which reduces the size of the matrix generated in the previous layer, and iii) The Smoothed layer smooths the output of the previous layer.

Stage 2: The objects extracted by CNN are introduced into the RNN model. It is noteworthy that the RNN network uses a gating mechanism that allows you to selectively save or forget information

about previous time steps, effectively fixing long-term dependencies in the input sequence. Finally, the RNN network output is transmitted to a fully connected layer to create dynamic forecasts based on time series data.

5.1. Convolutional neural networks

Convolutional Neural Networks (CNN) represent one of the most sought-after deep learning architectures and have demonstrated outstanding results in tasks such as image processing, computer vision, and pattern recognition [11]. Their effectiveness is due to the use of a key mathematical operation — convolution, which underlies the extraction of features from the input data.

From a mathematical point of view, convolution (within the framework of functional analysis) is an operation on two functions, as a result of which a third function is formed, describing how one of the original functions changes the shape of the other. This operation can be implemented either as an integration or as a discrete summation, depending on the nature of the data being processed. Convolution has a wide range of applications, from statistics and digital signal processing to numerical solution of differential equations, computer vision, engineering modeling and, of course, machine learning [11].

In the context of neural networks, the convolution operation makes it possible to effectively identify local patterns and spatial dependencies, which makes CNNs especially useful in analyzing structured data, including images, video streams, and even representations of network traffic in information security tasks.

Mathematically, convolution is defined as follows (one-dimensional example):

Let's define a continuous function $y(t)$ given by the formula:

$$y(t)=\int x(a)\omega(t-a)da, \tag{1}$$

where x(a) represents the so-called input data, and $\omega(t - a)$ is usually called the weight function or kernel.

The above integral is written in a more compact form as:

$$y(t)=(x\cdot\omega)(t), \tag{2}$$

The discretized version says:

$$y(t)=\sum_{a=-\infty}^{a=\infty} x(a)\omega(t-a). \tag{3}$$

The calculation that is the reverse of the convolution operations described above is known as deconvolution. It is used to reconstruct the original input signal $x(a)$ from the known output signal $y(t)$ and the impulse response of the system $\omega(t)$. In other words, deconvolution allows you to restore the original signal, distorted when passing through a system with a known or unknown impulse response. In cases where several signals overlap, deconvolution helps to separate them for later analysis [9].

Thus, deconvolution is a powerful tool in various scientific and engineering fields, allowing the recovery and analysis of signals and images distorted during transmission or recording [12].

5.2. Recurrent neural networks

Recurrent Neural Networks (RNN) play a key role in sequential information processing and are widely used in time series forecasting, speech recognition, and machine translation. Unlike direct

propagation neural networks, where data is transmitted only in one direction – from input to output, RNNs have internal feedbacks, which allows them to take into account the context of previous states [10].

Due to their architecture, RNNs are capable of processing arbitrary-length inputs, including text sequences, audio signals, and time series. This makes them particularly effective in tasks where time dependencies need to be taken into account. Their ability to remember information about previous steps and update the hidden state using nonlinear dynamics provides high flexibility and expressiveness when modeling complex processes.

The standard RNN processes sequential data by maintaining the hidden state ht at each time step t based on the input data xt and the previous hidden state $h_t - 1$. The latent state update equation has the form:

$$h_t = f\left(W_h h_{t-1} + W_x x_t + b\right), \tag{4}$$

where, $W_h \in R^{n \times n}$is the weight matrix, $W_x \in R^{n \times m}$ is the input weight matrix, $b \in R^n$ is the offset, and f is the nonlinear activation function.

The output of y_t at each time step is calculated by applying a linear transformation to h_t:

$$y_t = g\left(W_y h_t + c\right), \tag{5}$$

where, $W_y -$¿is the matrix of output weights, c is the output offset, and g is the softmax function for classification tasks.

This formula describes the last step of direct propagation in the model: the transformation of the latent state ht into the output vector y_t, where each element represents the probability that the input example belongs to a certain class.

In tasks such as text tonality analysis or document classification, the hidden states of recurrent neural networks (RNNs) are transformed into class probabilities using this formula. This formula is used on the last layer to predict the probabilities of an image belonging to different classes, and in speech recognition systems, the output of the model is transformed using softmax to determine the probability of each possible word or sound [13].

6. Conclusion

The conducted research was aimed at verifying and evaluating the effectiveness of the signal indicator control algorithm. The analysis showed that the algorithm proposed by the authors provides a clear and deterministic logic for processing input control signals and generating the corresponding output signal state.

The developed algorithm is a fundamental and verifiable component of the signal indicator control system that meets high requirements for safety and determinism and demonstrates an example of effective programming of finite automata for railway signaling systems. It confirms the possibility of implementing complex logical dependencies using standard boolean operations, which simplifies verification, increases system reliability, and facilitates further certification.

The conducted research has confirmed the relevance and effectiveness of using deep learning methods, in particular convolutional (CNN) and recurrent (RNN) neural networks, to increase information security in the context of digitalization of railway infrastructure. These architectures allow you to take into account both spatial and temporal characteristics of data, ensuring high accuracy in detecting information threats.

Self-learning models should also be highlighted as a promising tool for analyzing unlabeled data. Their key advantage lies in their ability to adapt to previously unknown attacks without the need for constant manual marking or specialist intervention. This is especially true in the context of a rapidly evolving cyber threat in critical infrastructure.

Automatic encoders, recurrent and convolutional neural networks remain relevant and can be effectively used as independent solutions for individual tasks (for example, anomaly detection or video stream analysis), as well as as parts of more complex hybrid systems.

Thus, deep learning is a powerful tool in the arsenal of information security tools for railway transport, which can significantly increase the industry's resilience to modern threats.

Further work will focus on refining the CNN-RNN model, exploring hybrid artificial intelligence approaches, and conducting pilot implementations to test its applicability in the railway industry. This will make it possible to create a more stable, intellectually secure information security system at railway infrastructure facilities.

This work is the result solely of the intellectual activity of the authors. Generative AI was not involved in the development of methodology, data collection or analysis, as well as in the formation of scientific conclusions.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] K. Sansyzbay, Y. Bakhtiyarova, T. Iliev, L. Tasbolatova, D. Sagmedinov, Method of Evaluation of the TETRA Standard Data Transmission Channel for Ensuring Information Security of the Railway Transport System. TEM Journal, 13(3), (2024) 2512-2521. doi:10.18421/TEM133-77.

[2] K. M. Sansyzbay, A. A. Kuandykov, Ye. A. Bakhtiyarova, S. V. Vlasenko, O. Zh. Mamyrbaev, Radio communication channel interaction method, maintaining train performance information security. Journal of Theoretical and Applied Information Technology, 98(06), (2020) 957-969.

[3] L. Tasbolatova, K. Sansyzbay, M. Orynbet, U. Maschek, Ye. Bakhtiyarova, A. Turayeva Development of a Radio Data Transmission Channel Model to Ensure Train Safety, Journal of Internet Services and Information Security (JISIS), 15(2), (2025)622-640. doi: 10.58346/JISIS.2025.I2.043.

[4] P. D. Mylnikov, P. A. Popov, Information Security in European Train Control Systems for Railway Transport. Intellectual Technologies on Transport, 3, (2016) 50-55.

[5] K. Sansyzbay, Y. Bakhtiyarova, T. Iliev, G. Patokin, L. Tasbolatova, D. Sagmedinov, Development of an Algorithm for a National Microprocessor-Based Centralization System With a Modular Architecture KZ-MPC-MA Featuring Advanced Intelligent Control Functions. IEEE Access, 12, (2024) 193229-193240, doi:10.1109/ACCESS.2024.3521219.

[6] K. Sansyzbay, E. Bakhtiyarova, T. Chigambaev, O. Abdirashev, Z. Badanbekkyzy Digital national microprocessor-based semi-automatic interlocking system to increase railroad capacity in Kazakhstan. Bulletin of L.N. Gumilyov Eurasian National University Technical Science and Technology Series, 150(1), (2025) 148–161. doi:10.32523/2616-7263-2025-150-1-148-161.

[7] K. M. Sansyzbay, G. S. Patokin, M. S. Batyrhanov, Y. A. Bakhtiyarova, D. B. Sagedinov National microprocessor centralization system with modular architecture KZ-MPC-MA,'' Patent Republic Kazakhstan, 36 788, Jun. 14, 2024.

[8] C. Livier Recio-Colmenares, R. Berenice Recio-Colmenares, R. Fernando Conchas-Cedano, I. Pilatowsky-Figueroa, C. Alejandro García García, Data-Driven Hybrid CNN-LSTM Neural Networks for Predicting Drying Kinetics of Green Tomatoes (Physalis ixocarpa) Integrating Mathematical Models, IEEE Access, vol. 13, (2025) 57413-57425. doi: 10.1109/ACCESS.2025.3554492.

[9] Weiqi Yan Convolutional Neural Networks and Recurrent Neural Networks, In book: Computational Methods for Deep Learning, (2023) 69-124. doi:10.1007/978-981-99-4823-9_3.

[10] M. Noguer I Alonso The Mathematics of Recurrent Neural Networks, (October 27, 2024). Available at SSRN: http://dx.doi.org/10.2139/ssrn.5001243.

[11] K. Oh, M.Yoo, N. Jin, J. Ko, J. Seo, H. Joo, M. Ko, A Review of Deep Learning Applications for Railway Safety, Appl. Sci. 12, (2022) 10572. doi: 10.3390/app122010572.
[12] M. Di Summa, M. Elena Griseta, N. Mosca, C. Patruno, M. Nitti, V. Renò, E. Stella A Review on Deep Learning Techniques for Railway Infrastructure Monitoring, IEEE Access, vol. 11, (2023) 114638-114661. doi: 10.1109/ACCESS.2023.3309814.
[13] P. López-Aguilar, E. Batista, A. Martínez-Ballesté, A. Solanas Information Security and Privacy in Railway Transportation: A Systematic Review, Sensors 22, (2022) 7698. doi: 10.3390/s22207698.