DEEP NEURAL NETWORK FOR ANOMALY DETECTION IN ACCELERATORS

M. Piekarski, W. Kitka, J. Jaworek-Korjakowska

National Synchrotron Radiation Centre SOLARIS, Jagiellonian University, Czerwone Maki 98, 30-387 Kraków, Poland

ABSTRACT

The main goal of NSRC SOLARIS is to provide scientific community with high quality synchrotron light. In order to do this it is essential to monitor subsystems that are responsible for beam stability. In this paper a deep neural network for anomaly detection in time series data is proposed. Base model is a pre-trained, 19-layer convolutional neural net-work VGG-19. Its task is to identify abnormal status of sensors in certain time step. Each time window is a square matrix so can be treated as an image. Any kind of anomalies in synchrotron's subsystems may lead to beam loss, affect experiments and in extreme cases can cause damage of the infrastructure, therefore when anomaly is detected operator should receive a warning about possible instability

ANOMALY DETECTION SOLUTION

The main purpose of the proposed solution is the detection of anomalies in the signals coming from the measuring devices in the SOLARIS synchrotron and as a result the detection of beam instability. Due to the complexity of the synchrotron operation and the number of subsystems necessary for its proper operation, it has a distributed control system. Measurement data from most devices are read and archived in database. One such key subsystem is the vacuum subsystem. The base concept of the proposed system is the fact that the collected measurement data from many devices at once and divided into appropriate time windows can be treated as an image (matrix of values). Therefore, the use of available algorithms and models for the purposes of classification of aggregated measurement data is authorized. In particular, this applies to deep, convolutional neural networks, whose ability to extract complex image features can help in finding those in multivariate signals. Proposed model is based on the pre-trained, deep, convolutional neural network VGG-19. Its biggest advantage is generalization for a wide range of data sets and no need to train all hidden layers. The system operation diagram has been presented in figure 1.

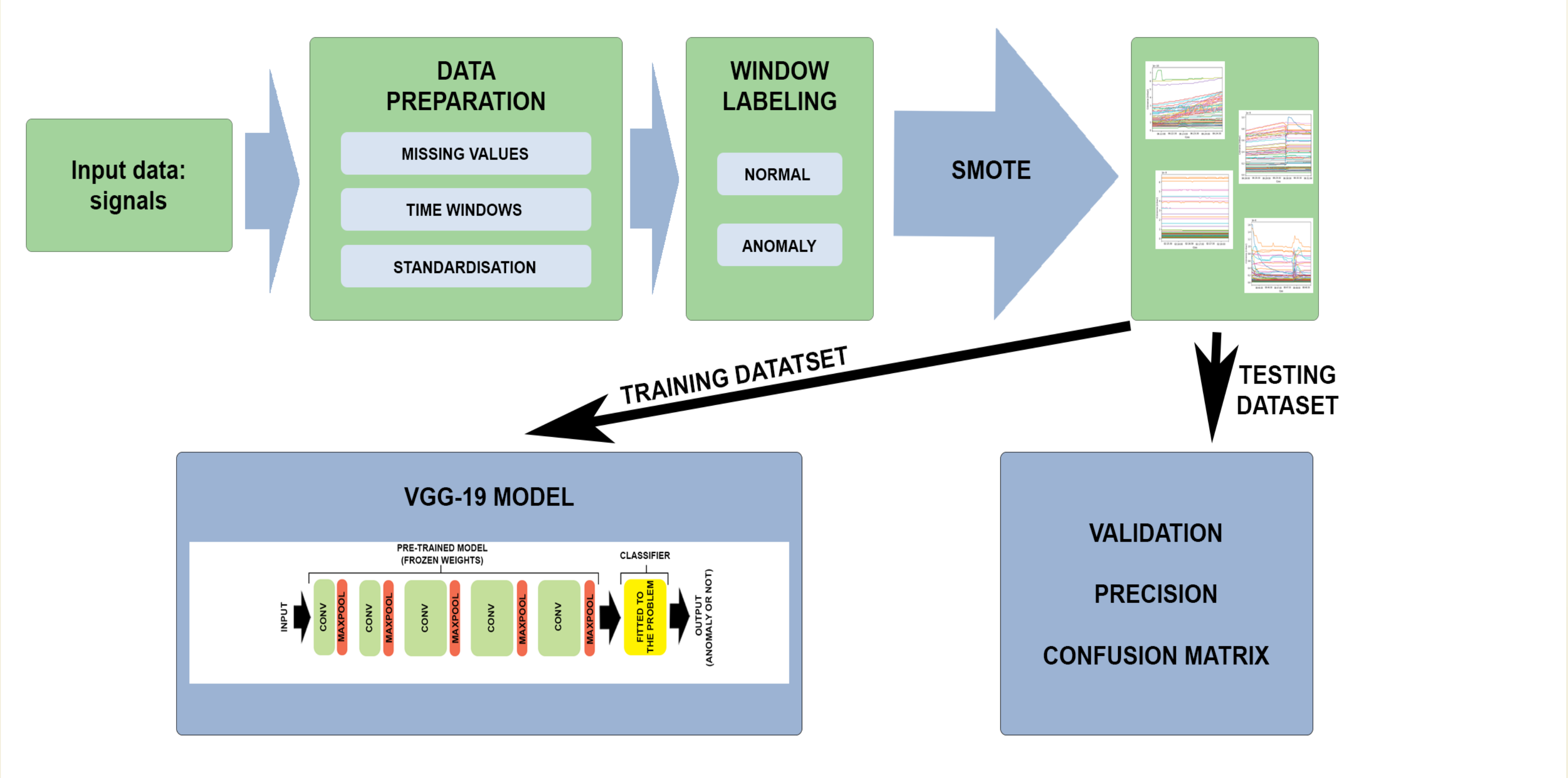

Fig. 1 System diagram

One of the most important stages of work on a system using machine learning or neural networks is proper data preparation. This allows to obtain satisfactory results of the model and also to avoid errors caused by entering the wrong or noisy data into the system. The data are pressures measured by 64 Gamma Vacuum power supplies located around the storage ring. Main part of data processing was to divide the collected measurement signals into identical time windows. Because there were 64 readings, the window size was assumed for 64 subsequent data samples. Thanks to this, square images were obtained, which can be easily used as input to the VGG-19 model. Due to the use of the supervised approach, each window prepared in this way had to be described as correct or anomaly. The assessment itself was done manually. Next step was data standardization. Example time windows are shown in figure 2.

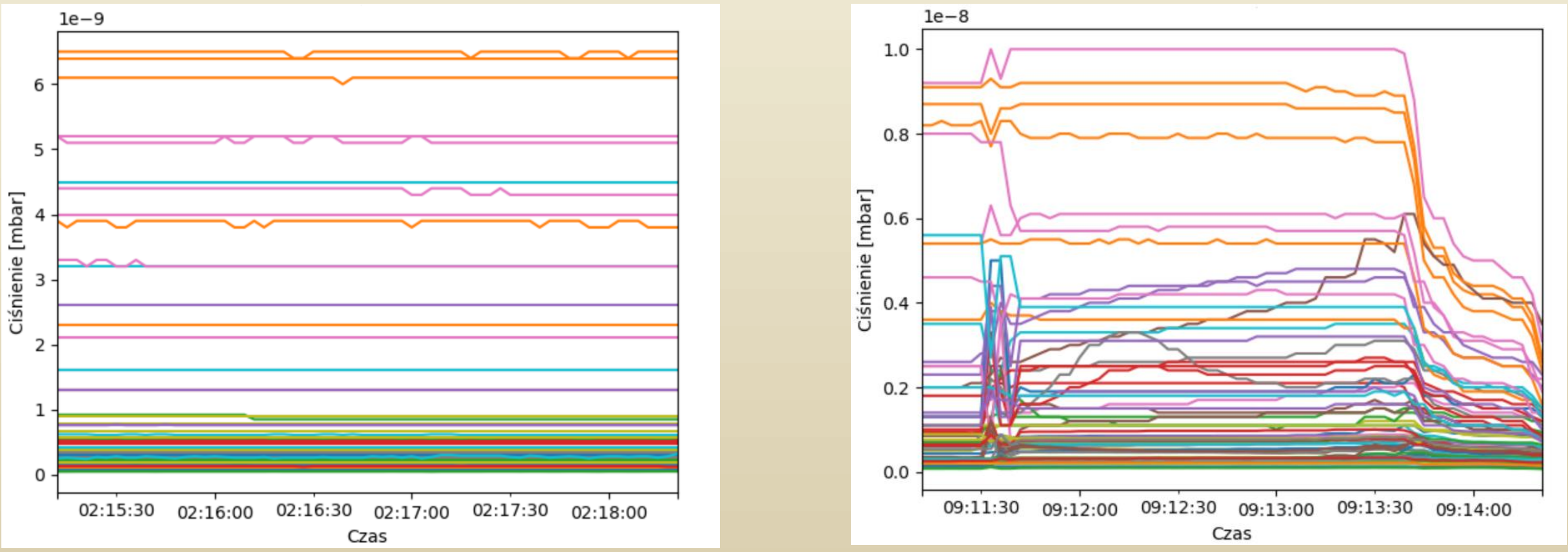

Fig. 2 Time window without and with anomaly

One of the difficulties encountered during the construction of the anomaly detection system was the data set imbalance between the anomalies and non-anomalies. Using such data to teach the model could result in incorrect operation of the system, which would be insensitive to windows that are anomalies and would give erroneous, high results. One of the techniques to solve the problem of unbalanced classes is the SMOTE algorithm (Synthetic Minority Over-sampling Technique) which works by generating new, synthetic data. In the proposed system, the classifier was created specifically for the problem of recognizing two classes: anomaly and correct signal. It consists of three layers. The first is the Fully-Connected layer with ReLU activation function, the next is the Dropout layer and the classifier finishes the Fully-Connected output layer with two outputs and the Softmax activation function. At each of these outputs, the network calculates the probability that the input image is an anomaly or correct signal.

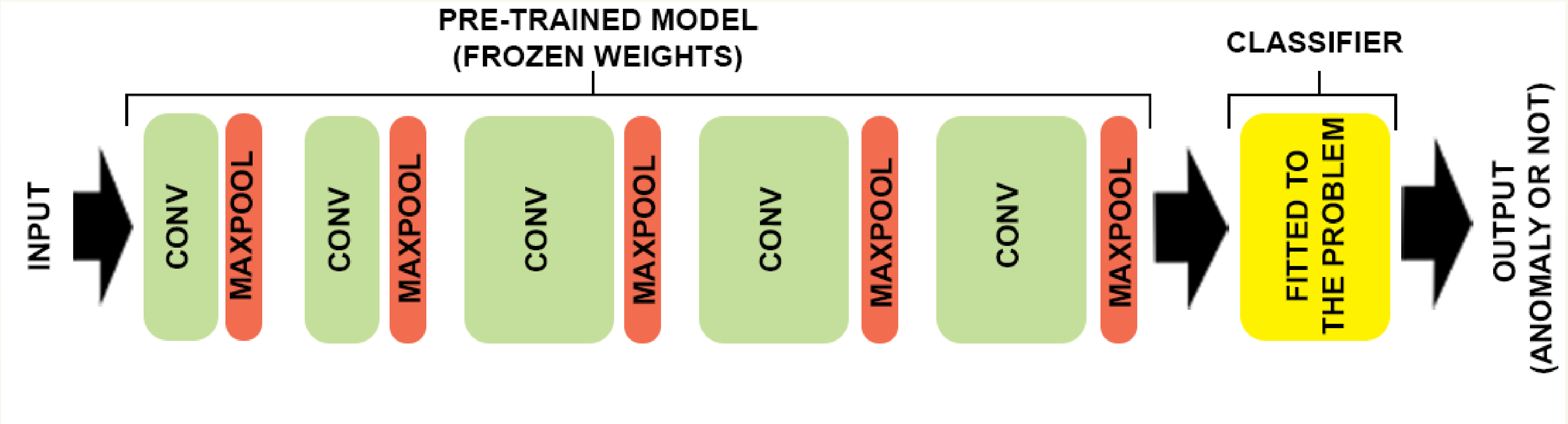

Fig. 3 Proposed architecture

TESTS AND RESULTS

At the stage of preparation and initial data processing, two data sets were obtained: training and test. They come from completely different periods of time, which ensures the reliability of the results obtained in tests. In both sets, the classes have been balanced using SMOTE algorithm. In addition, for the purposes of optimizing model parameters the validation set was created. Results are presented in table 1. Very high results of the proposed model were achieved. Based on these results, it can also be concluded that the system was not overfitted. For further evaluation of the model, confusion matrix was created (see table 2). The system should have as many classifications as possible on the diagonal of the confusion matrix (correct classification), and when there is a mistake it should be FP (false positive) type. It means that it is a false alarm which can be easily verified by duty operator.

Tab. 1 Accuracy of the model

	Data sets		
	Training	**Validation**	**Test**
Accuracy	96.70%	98.35%	95.13%

Tab. 2 Confusion matrix

Actual class	**Predicted class**	
	positive	**negative**
positive	820	2
negative	78	744

Figure 4 presents the classification of time windows by the proposed anomaly detection system applied to the original signals. The time of detection of the anomaly is closely related to the construction of the time window. The time window of 64 samples is about three minutes so this is the maximum time after which the system should detect the anomaly. Fragments qualified as positive results are marked in red.

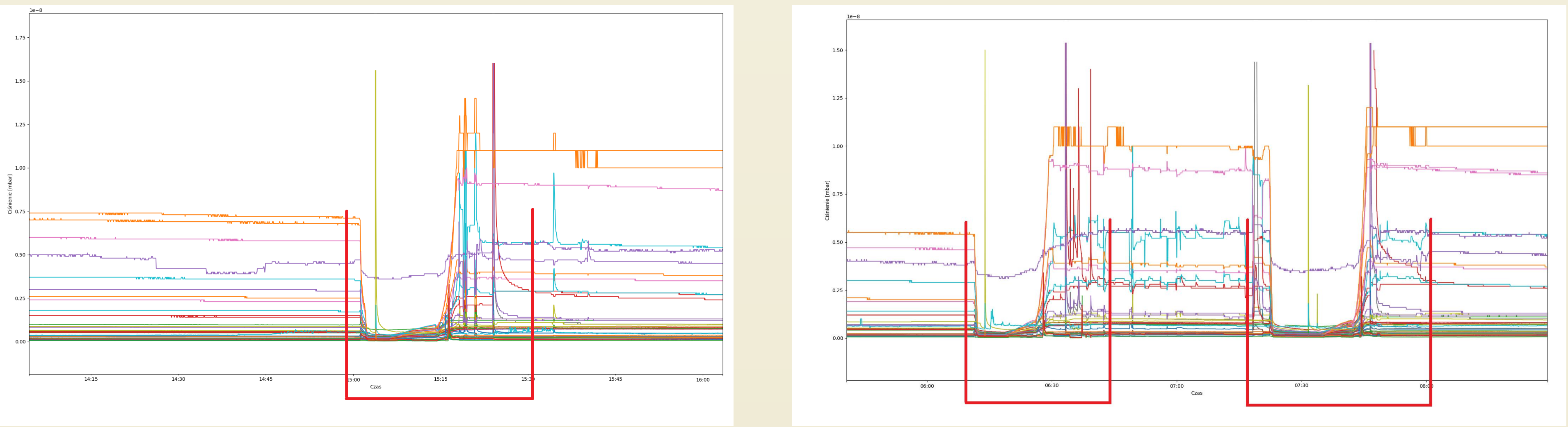

Fig. 4 Time windows with detected anomalies

CONCLUSION AND FUTURE WORK

The paper presents the design of the anomaly detection system in the multivariate signals of the SOLARIS synchrotron. It was inspired by earlier research conducted using various machine learning techniques to predict and classify anomalies and potentially dangerous situations in accelerators. The proposed solution is based on the deep, convolutional neural network VGG-19. It is a pre-trained network with optimized weights, with high generalization capabilities. It has been shown that a set of measurements in a certain time window can be treated as an image. The VGG-19 network with data prepared in this way allows detecting anomalies in signals. In addition, it was pointed out that the possibilities of adapting architecture to own needs by using transfer learning are so large that the class of problems to which VGG-19 can be used as the base model is very wide. The system has great development potential. The next stage will be operation in real time so aggregation of the measurements into time windows and their classification. Further modifications may include extending the number of classes in such a way as to include the SOLARIS' state machine and by that better adapt to the specifics of the machine operation.