Technological aspects of deep-learning algorithm development for processing information in fibre-optic trunk pipeline security systems

by A.V. Makarenko

Constructive Cybernetics Research Group Moscow, Russian Federation and Institute of Control Sciences (Russian Academy of Sciences), Moscow, Russian Federation

Article submitted for publication by OMEGA JSC, Moscow, Russian Federation

This paper examines the key issues in constructing algorithms for processing target information in fibre-optic trunk pipeline security systems. Deep learning methods are applied in order to implement the basic functions of the environmental recognition system in the pipeline area, in particular: detecting and classifying input signals, extracting signal-event tracks, and identifying events and their sources. This study also presents a plan for walkthrough development of algorithms for processing information based on deep-learning methods, from task assignment to reference implementation and prototype testing. The basic stages of the plan are illustrated by studies which were carried out during the development of algorithms for primary signal classification in the leak detection and activity monitoring system on behalf of AO Omega, Moscow.

Key words: deep learning, signal detection, classification, convolutional neural network, recurrent neural network, fibre-optic system, trunk pipeline security.

Distributed fibre-optic vibration sensors, which operate on the basis of the reverse-Rayleigh scattering effect, are currently one of the most promising solutions in trunk pipeline security. Among the considerable advantages of this solution are the following:

• The system can be hidden, and the fibre-optic sensor is totally inactive (a fully dielectric version is also possible). For system unification, a standard telecommunication fibre-optic cable can be used as a sensor.

• Security boundaries up to 100 km long can be covered with one electronic block. One such
block can replace many individual sensors and is more technically and economically viable in security systems for long-distance perimeters.

- The system’s spatial-resolution capability is high, at no lower than 5 m, and uniform along the whole length of the security boundary. The fibre-optic sensor is highly sensitive to longitudinal strain in its laying surroundings – from 70 nanometres.

- It is possible to detect the direction from which trespassers cross the security perimeter (using several spatially distributed parallel fibre-optic lines), as well as the locations of trespassers within the security perimeter (using specialized layouts of the fibre-optic cable).

- Practically any event can be detected and classified (as it is accompanied by geo-, hydro-acoustic, or seismic waves) when it occurs at the security perimeter (close to the fibre-optic sensor).

The operating principle for the class of security systems described here is based on the timed detection of reverse-Rayleigh scattering of a short laser pulse moving along the optic fibre [1]. In a coherent reflectometer (see Fig.1 for a general structural-functional scheme), a highly-stable narrow-band reference laser is used. From this, the intensities of reverse scattering signals are totalled, taking into account phase delays. In turn, oscillations in the sensor’s surroundings cause micro-vibrations to the optic fibre in the sensor and displace the Rayleigh scattering centres relative to one another. In this way, the spatially localized disturbances to the optic sensor cause local modulation to the intensity of the coherent reflectogram.

Thus, based on an analysis of temporal, spatial, and frequency characteristics of local modulations in intensity for the coherent reflectogram, it is possible to build algorithms for detecting and classifying the sources of disturbances in the vibration-acoustic field close to the security zone of the facilities being monitored [2].

In view of the device’s operating principle and a range of limitations caused by the current level of technical development, building algorithms for detecting and classifying events in the security zone of a trunk pipeline is complicated by a range of factors. It requires the use of special technology to synthesize high-quality solutions, which are suitable for industrial application. The present article is dedicated to one approach to solving this problem, which was developed by the author and is based on deep-learning technology.

**Algorithms for detecting and classifying events in fibre-optic perimeter security systems**

**General provisions**

The process of detecting and classifying input signals in fibre-optic perimeter security systems is shown in Fig.2 in a general view in the classical form [3].

Readouts of input (observed) signal are input to the matrix:

\[
I_{k',l'} = \begin{bmatrix}
I_{k'=K',l=1/2} & \cdots & I_{k'=K',l=1/2} \\
\vdots & \ddots & \vdots \\
I_{k'=L',l=1/2} & \cdots & I_{k'=L',l=1/2}
\end{bmatrix}
\]

where:

- \(K'\gg1\) given that signal detection procedures from a single readout are, as a rule, variable [3];
- \(L'\gg1\), based on conditions for detecting (selection) moving and/or spatially extended signal sources.

Further, the data block \(I_{k',l'}\) is input to detection procedures. Based on their output, evaluations of \(a\ posteriori\) probabilities are formed:

\[
p_{k',l'}^{w} \text{ only 'background' is present (natural and/or industry-related interference)}
\]
Besides interference, there is also an ‘event’ subject to monitoring.

It must be noted that these two probabilities form a full group \( p_s' \), \( p_s'' \). Where the probability \( p_{SE,k',p}^{ps} + p_{SG,k',p}^{ps} = 1 \) exceeds the threshold value \( a_d \), the classification procedure then comes into play. From its output, evaluations of a posteriori probabilities \( p_{m,k',p}^{ps} \) are formed. The ‘event’ detected is attributed to one or another previously defined class, where \( m = 1, M; \ M << \infty \) is the number of recognizable classes. As a rule, the aim of classification is set as single-class, i.e.

\[
\sum_m p_{m,k',p}^{ps} = 1
\]

and

\[
\epsilon_{k',p}^{ps} = \arg \max_m p_{m,k',p}^{ps} .
\]

As has already been indicated, developing algorithms for detecting and classifying events in fibre-optic security systems for trunk pipelines is complicated by a range of factors. The fundamental problems include:

- Deviation of the local temperature of the sensitive element of the sensor causes incidental modulation to the reflectogram.
- Phase (frequency) deviation in the laser radiation causes local areas of temporary insensitivity to spontaneously appear and disappear in the sensor.
- A significant change to the signal/noise ratio along the length of the optical sensor, owing to the natural process of attenuation in the probing laser pulse and scattered response with their propagation along the optic fibre.
- The coherent reflectometer in the so-called ‘non-linear’ mode loses information about the direction (compression/extension) of longitudinal strain in the optic fibre. Moreover, it has no single consistent scale for measuring the intensity of the input impacts and the level of internal noise.
- The large variation and instability of potentially possible signal portraits for recognizable events requires special approaches to formulating dictionaries for identifying events.
- In real settings, as a rule, there are
intensive signal-like interferences.

- The wide variation and instability of mechanic-acoustic properties in the sensor laying surroundings lead to substantial change in the structural (frequency) characteristics of signal portraits of recognizable events.

As a whole, these factors form a highly complex signal-interference environment. Algorithms for recognizable events must function in these conditions. A further complicating factor is the highly rigorous requirements, customary in monitoring and security systems for important facilities, for the level of errors of the first and second type when making decisions linked with detecting and classifying events in a protected zone. Consequently, event-recognition algorithms must possess a significant depth of adaptation in order to function well and reliably.

Target information processing circuit

The direct application of algorithms for detecting and classifying events in fibre-optic perimeter security systems in the form shown in Fig.1 is significantly complicated for technical reasons. These complications can be examined through a concrete (and quite typical) example.

For the optical pipeline-monitoring system (OSMT) which has been developed by Omega Co, the extent of the monitored area (one electronic block) consists of 100 km, space step of the sensor is 2 m. The sampling frequency (for one spatial channel) is 1,666 Hz, and the ADC (analogue-to-digital converter) capacity is 16 bits. As such, the reliable detection and classification of, for example, an event such as a leak of liquid from a pipeline into the soil requires a value of $K' > 1.8 \times 10^8$ (no less than 30 minutes of physical time for a leak with pressure of 8 at. and hole diameter of 13 mm), which automatically puts a limit on the server's RAM capacity of no less than 1 TB. At the current level of computer technology development, this requirement is economically unjustifiable. Therefore, a different application is proposed for the target information processing circuit in fibre-optic perimeter security systems as part of events detection and classification (see Fig.3) [4]. There are several key differences in comparison with the classic plan.

Firstly, the detection and classification procedures are combined into a single operation: primary classification of input signals at the level of spatial-temporal blocks $I_{k',l'}$. As such, the output is

$$\sum_m p_{m, k', l'} = 1,$$

where $m = 0, M$, but $m = 0$ is the 'background' class. It must be noted that in this case the value of $K'$ is determined not by accumulation time for reliable signal detection, but by the duration of the informational pattern of the recognized signal.

$$P_{s, k', l'}^{ps} = \begin{bmatrix} p_{s, k' = K'^2, l' = L'^2} & \cdots & p_{s, k' = K'^2, l' = L'^2} \\ \vdots & \ddots & \vdots \\ p_{s, k' = K'^2, l' = L'^2} & \cdots & p_{s, k' = K'^2, l' = L'^2} \end{bmatrix}$$

(2)
Secondly, in order to reduce the probability of omitting events and to stabilize the rate at which false alarms are generated, the primary classifier decisions at the level of separate spatial-temporal blocks are integrated (together with the relevant filtration) into signal-event tracks (see Equn 2 below, left), where $K'' > 1$ and $L'' > 1$. On this basis, final event hypotheses can then be formed.

Thirdly, the secondary classifier has a range of special features:

- the operator can manage its decisions by setting the allowable levels of first ($\alpha$) and second ($\beta$) type error.

- additional information $V_{k''}^{a''}$ is input into this classifier, namely: a priori probabilities of the presence of a particular event class at a particular time in a particular spatial field; data from the output of the primary classifier of the second device (where there is opposite parallel monitoring of the security perimeter by two devices); the decisions of additional monitoring systems, which function using different physical principles.

- Thus, a two-level scheme for implementing the target information processing circuit in fibre-optic systems for perimeter protection in the area of detecting and classifying events (see Fig.3) not only saves considerable amounts of RAM in the computational block, but also allows better-quality narrowly-specialized algorithms to be synthesised.

**Primary classifier of input signals**

Various approaches may be used to build the primary classifier of input signals, which are embedded in the target information processing circuit. Firstly, there are classical methods of synthesizing signal-detection algorithms [3, 5]. As our research has shown, the use of these methods is highly ineffective given that the empirical probability densities for signal and interference qualitatively differ from the theoretical probability densities. The algorithms are synthesised on the basis of the latter. In turn, the use of non-parametric algorithms [3] is not effective due to significantly unstable empirical densities. Secondly, it is possible to implement the primary classifier based on classical, so-called shallow-learning methods of machine learning, such as SVM, Random Forest, logistic regression, etc. [6]. Although in practice these methods can be used successfully to identify events in fibre-optic perimeter security systems [2, 7], nonetheless our research has shown that implementing industrial-scale systems on the basis of classical methods of machine learning is not effective for a whole range of reasons. Among the most important are the following:

- Exclusively empirical (hand-engineered) synthesis of high-level informative features. This situation makes it significantly more complicated to change sets of recognizable events. Moreover, as a rule, similar hand-engineered lose effectiveness where there is a significant change to the parameters of the signal-interference setting. Furthermore, the process of classifier synthesis then includes the human factor, and the involvement of highly skilled subject specialists becomes necessary to hand-engineered extract information indicators from the data.

- Studying models of data with a great many variables is very complex, as such data have a hierarchical structure (often unknown a priori) of a descriptor along time and/or spatial axes, including due to the linearity of many classical algorithms for classification [6].

Deep-learning algorithms based on deep neural networks were therefore used in
order to synthesise the primary classifier [8]. These algorithms possess two significant advantages in the context of the task under discussion:

- This class of machine-learning algorithm attempts to model hierarchical abstractions in data, using architectures consisting of a cascade set of non-linear transformations (filters), which immediately increases the generalization ability of the model [6, 8].

- This class of machine-learning algorithm works with raw data (low-level features) and independently extracts (generates) an indicator description of the facilities [6, 8]. This is, in other words, meta-learning – the algorithm can train itself independently how it can learn better. Such algorithms can therefore – in a very real sense – be called intelligent.

As a result, the primary classifier for input signals was implemented in accordance with the structural-functional diagram presented in Fig. 4.

A preliminary filter is necessary to delete components lying outside the frequency range for useful signals. The presence of this filter is absolutely necessary when the time-domain decisive statistics are used. The look informed is implemented according to Equn 1. For each data block the decisive statistics are calculated both in the time domain (coefficients of kurtosis, skewness, etc.) and in the frequency domain (filter bank [9], etc). As decisive statistics are formulated, a secondary filter is included in the circuit so that the recognition algorithms function with better stability and quality. Given its high \textit{a priori} uncertainty, the secondary filter is built as a blind processing type [10]. Next, an array of primary indicators is formed from the set of decisive statistics. This array has a special structure, co-ordinated with the input layer of the neural network NN. Before being fed into the NN input, the array is transformed according to the special rules of the TF and normalized by the NF.

The general structure of the deep neural network, a key element of the primary classifier, is presented in Fig. 5.

As can be seen in Fig. 5, a tensor $\mathbf{T}_{k',f'}$ is applied to the input of the neural network. It has dimensions $K_{M} \times L_{M} \times R_{M}$ where $K_{M} > 1$, $R_{M} > 1$, and $L_{M} > 1$, i.e. the neural network processes the input data in co-ordination for several spatial channels, which allows moving and/or spatially extended signal sources to be detected (selected). It should be noted that the implemented condition $L_{M} > 1$ is a key difference in this approach when compared to one previously put forward [4]. The neural network is based on a convolution architecture. One argument in favour of this solution is its phenomenal potential for application in challenges based on the similar ideas [8, 9]. The output layer of the neural network contains $M + 1$ neurones (according to the number of recognizable signal classes).
and the SoftMax activation function is used in it:

$$p^s = \frac{e^{s_i}}{\sum_{j=0}^{M} e^{s_j}}, i = 0, M,$$

(3)

which allows values of the output vector $p^s$ to be interpreted as the events’ probabilities, whose population forms a full group [6]. For example, within the framework of OSMT studies (the developer Omega) the following classes were assigned for detected and classified events (signals):

- 0 - natural and technogenic interference
- 1 - liquid leak from the pipeline into the soil
- 2 - work with a manual digging tools
- 3 - work with heavy excavating machinery
- 4 - drilling the pipeline wall
- 5 - welding to the pipeline wall
- 6 - grinding the pipeline wall.

It should be noted that the signal (event) classes identified here in effect completely cover the basic monitoring objectives in the protected area of trunk pipelines and allow typical trespassing scenarios to be monitored.

The neural-network solution can be obtained on the basis of a fairly effective threshold rule, which the following adaptive properties can be fully attributed to Equin 4 (below) where $\alpha_s$ is the established vector of thresholds for decision-making for a neural network.

It is possible to make the given vector dependent on the spatial channel and the moment in time when the decision is made.

An example of the primary classifier functioning is presented in Fig.6. The diagram was formed after processing using rule (4) in validation sample #269 (this was not included in the train set).

$$c^s_{ak,l'} = \begin{cases} i^* \left( p^s_{ak,l'} \right)_r \geq (\alpha_s)_r, & \text{otherwise.} \\ 0 \end{cases}, \quad i^* = \arg \max_i p^s_{ak,l'}, \ i = 0, M,$$

(4)

It can be seen from Fig.6 that the primary classifier had no trouble extracting the spatial-temporal blocks, containing a class-6 test signal, from which the secondary classifier formulated a signal-event track. The parameters of the synthesized track matched reference parameters with high precision. This then meant a true-positive decision could be formulated about the presence of a class-6 event.

Detailed research should be carried out into the precision characteristics of the primary classifier and its components on the basis of confusion-matrix analysis [6], and also of the distribution densities for classifier response confidence under various conditions of signal-interference settings. An error matrix for the network trained within the framework of OSMT studies is presented as an example in Table 1.

Analysis of Table 1 shows that at this stage there is no point in having separate signal classes for drilling, welding, and grinding. It seems more logical to unify them in one integrated class ‘manipulation’, for all manipulations with the trunk pipeline wall. Quality indicators for this (unified) class then improve significantly: sensitivity = 74.49%; precision = 95.91%; F1 score = 83.86%.

It can also be seen from the data in Table 1 that the probabilities of errors received
at the output of the primary classifier absolutely must not be interpreted as final for event hypothesis, because the system also includes the secondary classifier (see the following section). This considerably alters the level of errors of the first and second type (and the correlation between them) during the transfer from solutions by blocks to solutions by tracks and the hypotheses as a whole.

In order to learn and test recognition algorithms, a reference library was formed for real signals (events) and interferences in conditions close to the normal mode of functioning for the device. The library was formed in three geographical zones, differentiated by soil properties, the way the fibre cable was laid, trunk pipeline parameters, and by the distinctive features of natural and/or industry-related interference. Recordings in all these zones cover several climatic seasons, which included the processes of freezing, thawing, and rainfall/flooding on the soils. It is also worth noting that in order to get a wide set of signal portraits for recognizable events, and in order to increase the classifiers’ generalizing capability, the leading parameters of signal sources varied significantly. For example, to record a class-1 signal (‘leak’) pipes of various diameter were used, and liquid pressure and fault size were changed. For class 5 (‘welding’) the electrode diameter and the welding current intensity were altered.

After the library was formed, 70 TB of data from it were initially formatted by the operator using automation means. Formatting included identifying the spatial-temporal borders of the signals, assigning the event class, and adding additional machine-readable meta-information. Next, the data were divided into learning and test samplings, which did not overlap. After initial learning and optimizing the classifier, the learning and testing data sets were formatted

![Diagram](image-url)

**Fig.6. The primary classifier response to validation sample #269, a class 6 signal, in which \( k_{\text{ref}}^b \) and \( k_{\text{ref}}^c \) are the time reference markers for the beginning and the end of the signal; \( l_{\text{ref}} \) is the reference (central) spatial channel; the grey dotted rectangle is the signal-event track of the signal, formed by the secondary classifier.**

**Table 1. Errors from the primary classifier (a test data set).**
more precisely by the classifier itself, after which the final tuning was carried out for the weights of the neural networks. The network was trained over about 200 epochs, using the SGD method [8]. The struggle with retraining was conducted by using the dropout technology and $L_2$ regularization. During training, the loss function was minimized by categorical cross-entropy:

$$L = {-1 \over D} \sum_{d=1}^{D} \sum_{i=0}^{M} (t^d_i) \ln(p^d_i),$$

(5)

where $D$ is the size of the learning set, $t^d_i$ is the vector of labels associated with a spatial-temporal block of data $d$, included into the training set. It can be noted here that in principle the target function during classifier training may be supplemented with a component which takes into account real (financial) losses and gains, in the event of particular classifier decisions.

In order to get additional gains, the architecture of the neural network (the sequence, the quantity and configuration of layers, and the quantity and configuration of convolution kernels, etc.) was optimized using genetic programming methods. The differential-evolution method was also used [11].

The solutions used here allow additional fine-tuning to be carried out for the primary classifier (the deep neural network) aiming to expand its adequate functioning area given the significant change in characteristics of the signal-interference setting. An effective mode of automatic additional learning (adaptation) for the deep neural network cannot be put into practice for a range of reasons, including the particular features of the system functioning as a whole:

There is a list in the retraining procedure (additional learning, adaptation) for the primary classifier’s deep neural network is highly cost-intensive from a computational point of view. It requires the use of a more-productive and expensive GPU (graphics’-processing unit) than in direct transit through the network. Installing such GPUs on every logical module of OSMT is not economically viable.

There is a list in the retraining process (additional learning, adaptation) for the primary classifier’s deep neural network requires high-quality learning and testing sets of data with considerable scope and variation in their characteristics. Moreover, the data sets are constantly changing and being modernized in the framework of adaptational procedures. This situation means that saving learning and testing sets of data in every logical module is not economically and organizationally viable.

As a conclusion to this section, it is worth noting that the median time spent by the GPU NVIDIA GeForce GTX Titan X on classifying packages from 1000 spatial channels of range is 62.278 ms per package (given mean-square deviation of 3.573 ms). Thus, in real-time mode the classifier allows decisions to be made throughout the monitored area at a range of 100 km, and still remains a computational reserve for further algorithm development.

The secondary classifier of event hypotheses

The primary classifier output (see for example Fig.6 and/or Table 1) is poorly applicable directly to forming notifications about the conditions which have arisen in the protected area of the monitoring system. As has already been
indicated above, in order to decrease the probability of omitting events and to stabilize the rate at which false alarms are produced, the primary classifier decisions must be integrated at the level of separate spatial-temporal blocks (together with the corresponding filtration) into the signal-event tracks. Based on these tracks, a final event hypothesis can then be formed. The above-mentioned points in our approach are covered by means of the secondary classifier. As can be concluded from Fig.3, the basic input data for the secondary classifier are tensors $P_{k,l}^j$, with dimensions $K_H \times L_H \times R_H$, where $K_H > 1$, $L_H > 1$, and $R_H = M + 1$. The primary and secondary classifiers can thus be stacked. A general structural-functional diagram for the secondary classifier of event hypotheses is presented in Fig.7.

Detailed examination of the organization of secondary data processing and synthesis methods is beyond the scope of the present publication, all the more so given that our research in these directions is actively continuing. It is worth mentioning, however, a few of the principal points which have already been achieved. Architecturally, the classifier consists of two blocks: TG – the signal-event track generator – is a deep recurrent neural network [8], which allows event characteristics such as causality to be brought into processing; the second block HG processes the parameters of the synthesized track and forms a decision as to the event hypothesis – the presence or absence in a given area of a particular monitored event.

Conclusion

In this study, we have shown that, on the basis of deep machine learning approaches, it is possible to create effective recognition algorithms for signals in distributed fibre-optic systems for monitoring and securing trunk pipelines. The studied neural network of the primary multi-class detector demonstrates - using real data - that the integral F1 measure of quality is no less than 83.86% of correctly recognized spatial-temporal data blocks which the primary detector operates. This assessment was obtained from test data (not included in the training set), in conditions as close as possible to real operating conditions for OSMT algorithms.

Unlike other approaches, as for example [3, 5], or with exclusively empirical (hand-engineered) synthesis of information features, deep machine-learning methods require less effort in development and, most importantly, they allow great flexibility for the system to be reconstructed given changes to the sets of recognizable events and/or given substantial change to parameters of signal-interference environment, i.e. they possess clearly expressed adaptive properties.

In conclusion, it is worth noting that the approach we present to developing recognition algorithms for signals and events on the basis of deep-learning methods has great potential. It could be successfully expanded into various connected fields, including: monitoring and management of trunk pipeline operating modes; ultrasonic or magnetic feature-detection in pipe walls; and technical infrastructure condition inspections. This conclusion was drawn based on the fact that the device we have been researching (the coherent optical reflectometer) functions in a significantly non-linear and unstable mode [1] in complex signal-interference conditions, but algorithms based on deep neural networks can counteract these negative factors with great success.

Bibliography

3. V.A.Bogdanovich and A.G.Vostretsov, The


