The Appliance of Deep Neural Networks in the Process of Managing Chemical Enterprises

E V Kulyasova¹, N S Kulyasov² and A Yu Puchkov²

¹Department of Economic Theory, D.Mendelev University of Chemical Technology of Russia, Miusskaya Square 9, Moscow, 125047, Russia
²Department of Management and Information Technology in Economics, National Research University «Moscow Power Engineering Institute» (Branch) in Smolensk, Energeticheskyy proyezd 1, Smolensk, 2014013, Russia

Abstract. This article is introduced into the perspective tendencies of the digital transformation of chemical enterprises which allow to improve the process of managing enterprises of the branch. Presented the algorithms of managing and technological information processing based on deep neural network apparatus. New approaches to data processing known as video analytics are applied; it allows to automate the registration process of visual data at chemical enterprises and reduce the impact of the human factor on the objectivity of the decisions. The worked out algorithmic structures of managing and technological information processing allows to carry out the identification of hidden regularities there and ensure the effective achievement of goals at chemical enterprises.

1. Introduction

Analyzing the perspective of the appliance of new digital technologies in industry, optimization effects of the implementation of innovative solutions in operating activities are firstly taken into consideration. Contemporary chemistry industry segments (mining, basic chemistry, organic synthesis, special chemistry) have a good marginality that is why top management is slow to apply innovative approaches towards technological equipment production. What is more, in their basis there is a well-known and reliable theoretical background while introduction of new technological cycles needs not only large material costs but also considerable effort to overcome bureaucratic barriers so typical for Russia.

However, as far as the introduction of advanced information technologies which ensure the basic industrial cycle is concerned the situation differs. Implementation of new solutions needs much less time and costs. The Government of the Russian Federation adopted the programme of development “The Digital Economics of the Russian Federation” in 2017 in order to avoid the technological backwardness of the country from the leading world economics and to implement the advanced technological solutions in informatization of all vital spheres of society.

According to the programme mentioned above, key technologies are big data, nanotechnology, and artificial intelligence, a system of distributed registry, quantum technology, new industry technologies, industrial Internet, robotics and sensor components, wireless communication technology, virtual and augmented reality technology. Analyzing the degree of their participation in different segments of Russian industry, it is necessary to note that chemical companies have remained unaffected by the innovation digital solution process.
Nowadays it takes within one-two years to get information from the chemical enterprises if it is possible to optimize production processes. Yet, even in this case, experts' expectations are not higher than 60% that the result will be optimistic.

2. Related works

Digitalization is one of components of creative industries. Digitalization implies penetration into all spheres of human activity of electronic infrastructure on the basis of development and use of modern information technologies. According to the Institute McKinsey by 2025 China will increase the country's GDP by 22% due to digital technologies, while the USA will earn 1.6 – 2.2 trillion dollars due to digitalization in added value of products and services. 60% reaches annual growth in the amount of data due to the increasing digitalization of the business.

The mentioned aspects contribute to the study of the possibilities of application of new information technologies in chemical industries. Taking in the consideration the fact that more than 70% of time research is spent on preliminary data processing, in other words, on routine work, it becomes obvious that solution of this task is better to transfer to the artificial intelligence. Taking into account the large volumes of technological data, their dynamics, variety of forms and structures typical for them, the use of intelligent methods of information processing is justified.

The peculiarity of intelligent data analysis methods is the possibility of their application in conditions of poor structuring of the original data and the solution itself. Fuzzy models that allow to aggregate the knowledge of experts and data of objective control over production processes into a single complex are well-established here [1]. A distinctive aspect of fuzzy logic models is their ability to solve standard optimization problems of production for the situation with incomplete data [2].

An important area of application of intelligent methods in production is machine vision, one of the ways of which is artificial neural networks. Among them, the most promising for image processing optimization are deep neural networks (Deep Neural Networks, DNN). Their successful application for image processing tasks has led to the appearance of a new direction of data analysis - video Analytics, which is a special variety of algorithms to find patterns in the information provided in the form of images. The relevance of research in this direction is justified by the obvious desire to exclude a person from another channel of information processing - the flow of visual data. This will significantly increase the amount of perceived and processed information in order to extract metadata on chemical processes and improve the reliability of management decisions.

A number of factors contributed to the development of the DNN device and its penetration into practical solutions:
- increased productivity of modern computing systems;
- successful technological solutions in the field of artificial neural network architecture used for pattern recognition;
- application of the neural network retraining procedure using a non-directional graphic model (restricted Boltzmann model);
- adaptation of graphics processors for neural network retraining cycles (first of all graphic processors of the company Nvidia);
- the appearance of a free tensor of processors Google.

The mentioned aspects enable project and apply deep neural network architectures which have rather high efficiency for computer vision and speech recognition tasks. Convolutional neural network (CNN) [3] has become the most popular recently. A distinctive feature of such networks from the previously known architectures is the limited fields of perception of the input image, the shared weights and the reduction of the dimension of the feature field. The basis of the network construction is the alternation of convolutional (C-layers) and sub-discretizing layers (S-layers). The output contains a fully connected layer (F-layers), the number of outputs of which determines the number of recognized classes of the input image. The alternation of C-layers and S-layers allows to map the characteristics of the cards of characteristics, which in practice means the ability to detect complex hierarchies of the input [4].
The proposed application of video Analytics in chemical production can complement the existing content of the control and measurement information of the enterprise about the ongoing processes and provide additional information for automated process control system.

3. Materials and methods
Among the most important areas of application of deep neural networks in the chemical industry in the framework of the program of digitalization of the economy can be identified the following tendencies:

1. Preprocessing of initial data to obtain the most important components necessary for further management and production operations.
2. Identification of objects and processes at their transition from one production zone or zone of responsibility to another according to the accepted technological and logical chains.
3. Automated control of the protected perimeter and compliance with production safety requirements.
4. The solution of the transport and logical tasks in terms of the high dimensionality of the original data, when linear programming techniques lead to a significant time-consuming to search for the optimal parameters.
5. Search and forecasting of faults and equipment failures, allowing to prevent emergency situations, which is especially important for chemical production.
6. Implementation of multi-dimensional machine vision based on the processing of information from a large number of sensors. This will allow to observe the course of chemical reactions and technological processes in real time in order to predict their behavior.
7. Business Analytics based on the generation of intelligent video cameras metadata about the state of the production process, as well as the extraction of hidden patterns in the visualized information about the results of commercial activities of the enterprise.

For all indicated areas, the overall structure of the video processing system coming from different sources will be the same (See figure 1).

Figure 1. The structure of the system of processing video data.
Information from video cameras installed in various places of the controlled process immediately enters the multiplexer. It alternately connects the information channels to the storyboard and analytical processing modules. The data coming from the sensors and from the external environment are pre-converted in the visualizer in the form of images, which then fall on the multiplexer. The analytical data processing module implements deep processing methods based on convolutional neural networks.

It is necessary to note that the expansion of analytical capabilities at the chemical enterprise is provided by the use of not only video cameras but also thermal imagers, radars, which are not sensitive to interference of the optical range. Such devices are already used in security systems of industrial enterprises. For example, radars and thermal images produced by AXIS company provide remote monitoring of the controlled area. Integration of data streams from these devices into the data of other information systems is carried out with the help of AXIS Camera Station, which allows you to create specialized intelligent applications for solving analytical problems.

The use of thermal imagers, video cameras, as well as a large number of other types of instrumentation is typical for the production of the chemical industry. Such equipment constitutes the lower levels of the hierarchical management system of the whole enterprise. If we talk about the higher levels implemented in the systems ERP (Enterprise Resource Planning), information from the external environment is added to these data, which also needs to be processed and taken into account when making management decisions.

The above mentioned DNN tendencies are based on the preliminary extraction of the required information from the images. This process involves different levels of video preprocessing complexity:

– defining the boundaries of an object is the lowest-level task;
– restoration of a three-dimensional image from a two-dimensional z by determining the vector to the normal;
– the definition of objects of attention, that is, parts of the image, which would draw the attention of a person;
– semantic segmentation that allows you to divide objects into classes by their structure, knowing nothing about these objects, that is, even before they are recognized;
– semantic highlighting of boundaries - a selection of the boundaries of recognized objects;
– selecting parts of the recognized object;
– the recognition of the objects themselves (the top-level task).

Depending on the purpose of application of neural networks in a chemical enterprise analytical data processing module provides solutions to problems of different levels on the basis of incoming video data.

The authors propose algorithms for detecting hidden patterns in visual displays of information presented in numerical form using CNN. The multichannel structure of the image processing algorithm from different sources provides the ability to take into account the importance of the channel information using weight coefficients \( w_i \) (see figure 2). The result of the algorithm is the vector \( Y \). It will contain one in the line, the number of which corresponds to the class number to which the object is assigned.

The algorithms are based on the unique transformation of the numerical source information into the image associated with the object and its further processing by CNN. The number of parameters simultaneously visualized on one associated image can be different, and the methods of this procedure will differ accordingly. Options images of different number of parameters \( X_1, X_2, X_3 \) are shown in figure 3. Figure 3.a bitmap showing the behavior of the X1 process over a period of time is shown. Setting the time interval allows you to take into account the dynamics of the process. Figure 3.b shows a visualization of the process described by two parameters, and figure 3.c – three parameters.

Image acquisition for one parameter \( X_1[t] \) is based on Fourier transform. Is the breakdown \( X_1[t] \) on K frames of N samples that overlap by N/2 of its width: \( X_1[t] \rightarrow X_1[n], n=1, \ldots K \). In each frame, a discrete Fourier transform is performed, the power spectral density [5] is calculated. Different frequencies can participate in further transformations with different significance which is given by different sensitivity. This is realized by means of a mel-scale, which sets the required sensitivity to frequencies. Mel-frequency cepstral coefficients are cepstrum values distributed over a Mel-scale
using a filter Bank. Then the matrix of cepstral coefficients is calculated and the image (bitmap) is compared to it. This image is used for recognition by the convolutional network because it is assumed that the image is unique [6].

![Diagram of the algorithm](image)

**Figure 2.** The structure of the algorithm of formation and processing of the image parameters of the chemical process.

![Diagram of types for displaying options](image)

**Figure 3.** Types for displaying options.

Each fragment was subjected to discrete Fourier transform and the power spectral density is calculated [7]:

\[
\text{Re}(S_n[k]) = \frac{2}{N} \sum_{i=1}^{N} X_{1,n}[i] \cos(2\pi k(i-1)/N),
\]

\[
\text{Im}(S_n[k]) = \frac{2}{N} \sum_{i=1}^{N} X_{1,n}[i] \sin(2\pi k(i-1)/N),
\]
where is k=1, ..., M, M=N/2.

The term cepstrum defines the result of a discrete cosine transformation from the logarithm of the signal amplitude spectrum [5]. On its basis, the Mel-frequency cepstral coefficients are calculated. When finding these coefficients a special Bank of filters is used, and in the simplest case, a uniform scale is used (the same throughout the frequency range). Calculations are carried out according to the following formulas and are reciprocal [8]:

$$f_{mel}(f_{hz}) = 1127 \ln(1 + \frac{f_{hz}}{700}), \quad f_{hz}(f_{mel}) = 700(e^{f_{mel}/1127} - 1),$$

where $f_{hz}$ is a frequency spectrum, $f_{mel}$ are mel-coefficients.

Find the lower $f_l$ and upper $f_h$ of the boundary frequency of the condition to ensure the correctness of the transformation; it is necessary that $f_h$ is not more than half the sample rate. To implement different sensitivity to frequencies, the number of filters $P$ in the filter Bank is set (in practice, P=12 is sufficient). Then frequency is transformed into mel:

$$f_{mel}^{m}(f_{l}), \quad f_{mel}^{m}(f_{h}).$$

Select the segment of mel-coefficients [$f_{mel}^{m}$, $f_{mel}^{m}$], divide into P+1 identical non-overlapping intervals [$f^{m}_{mel}$, $f^{m}_{mel}$], j=1,...,P+1 and calculate their middles: $C_m[i] = f^{m}_{mel} + i \text{ len}$, i=1, ..., P. After that, they are recalculated to the frequency scale (filter bandwidth centers) and the bandwidth centers (Hz) are transferred to the matrix count numbers $P_n[k]$:

$$C[i] = f_{hz}(C_m[i]), i = 1,...,P, \quad f_{amp}[i] = \frac{M}{F_s} C[i], i = 1,...,P,$$

where $F_s$ is the sampling frequency of the signal $S[t]$.

After calculating the sum of the products of the spectral density samples on the frequency response of the corresponding filter:

$$S_n[i] = \sum_{k=1}^{M} P_n[k] H[i], i = 1,...,P,$$

and conducting a cone transformation, we obtain:

$$C_n[j] = \sum_{k=1}^{M} \ln(S_n[k]) \cos(j(k - 0.5) \frac{\pi}{P}), i = 1,...,P, j = 1,...,J,$$

where $C_n[j]$ is a matrix of cepstral coefficients, J is a number of coefficients (J<P).

The obtained matrix $C_n[j]$ can be interpreted as an image (bitmap), reflecting the hidden characteristics and patterns of the analyzed suburban route. If necessary, you can set different sensitivity to frequencies using a nonlinear mel scale.

When analyzing two parameters $X_1$ and $X_2$, visualization can be performed on the basis of a simple point mapping in the Cartesian coordinate system. In the case of three parameters $X_1$, $X_2$ and $X_3$, the three-component system diagram method, often used in chemical studies, can be used [9].

Multiple images associated with an object or process can be mapped to the same object or process. Recognition of each associated image is performed by a separate CNN, each of which has the number
of outputs of the last, fully connected layer, coinciding with the number of classes. As a result of the work of the $i$ convolutional network, the vector of estimates of the object belonging to each of $a_1, a_2, \ldots, a_k$ classes is formed on the basis of the recognition $i$ associated image: $\text{ge}_i = \{\text{ge}_{i,1}, \text{ge}_{i,2}, \ldots, \text{ge}_{i,k}\}$, $i=1, \ldots, k$. By the number of classes, separate blocks of fuzzy logic output are created (see figure 2), and each $j$ block collects elements of GCI vectors from convolutional networks corresponding to the $j$ class at its input. The basis for the creation of fuzzy systems is the applied ontology of the subject area, which, we assume that it has already been developed. (This process deserves special attention, but does not have theoretical significance within the provided material.) Fuzzy inference blocks contain rule sets of the form:

$$\text{If } \bigvee_{i=1}^{kp_1} (\text{gc}_{i,1} \land \ldots \lor \text{gc}_{i,k} = T_{i_1} \land \ldots \lor \text{gc}_{i,k} = T_{i_m} \text{ TO } g_i = TG_j)$$

$$\text{If } \bigvee_{i=1}^{kp_2} (\text{gc}_{i,1} \land \ldots \lor \text{gc}_{i,k} = T_{i_1} \land \ldots \lor \text{gc}_{i,k} = T_{i_m} \text{ TO } g_i = TG_j)$$

$$\ldots$$

$$\text{If } \bigvee_{i=1}^{kp_k} (\text{gc}_{i,1} \land \ldots \lor \text{gc}_{i,k} = T_{i_1} \land \ldots \lor \text{gc}_{i,k} = T_{i_m} \text{ TO } g_i = TG_j),$$

where $kp_1, \ldots, kp_k$ is a number of rules for each class; $T_{i_1}, \ldots, T_{i_m}$ is a class of terms for the corresponding components of the vector of estimates of class membership $\{\text{gc}_{i,1}, \text{gc}_{i,2}, \ldots, \text{gc}_{i,k}\}$, $i=1, \ldots, k$, $TG_j$, $j=1, \ldots, k$ are the terms belonging to the classes of the components of the vector $\{g_1, g_2, \ldots, g_k\}$.

Sets of rules, their composition and content depend on the ontology of the subject area.

The program that implements the described method of fuzzy classification based on convolutional networks is written in Python 3.6 in the Spyder environment from the Python Assembly Anaconda, Linux operating system. To be able to use deep neural networks, TensorFlow machine learning library was additionally installed [10]. Note that CNN training is a resource-intensive process, as thousands of synaptic weights are adjusted on a large number of images and the use of the Central processing unit (CPU) leads to a long learning process. Therefore, it is advisable to transfer calculations to a graphics processor (GPU), containing thousands of microprocessors simpler than the CPU, but capable of performing many non-complex operations, paralleling the process of network training. Redirection computing libraries TensorFlow on the GPU is possible only for cards of Nvidia with support for CUDA-hardware-software architecture of parallel computing [11, 12].

The CNN architecture used in the various associated image processing channels is shown in figure 4.

When the program was running, the learning process was transferred to the GeForce GTX 1060 containing 1280 Maxwell/GP106 cores. It allowed to reduce the training time of the network more than 20 times compared with the training time on the CPU.

4. Results

The proposed algorithm of fuzzy classification was applied in the work of the HR-service of the chemical enterprise, numbering about one and a half thousand people of staff, for the analysis of working time loss. A significant number of members of the organization enabled to obtain a reasonable amount of training data for CNN.

The sources of the images were data from the electronic organization and photos of employees. The data from the entrance reflected daily (during the quarter) deviations from the standard time of passage of the turnstile at the entrance and exit. This data was then converted to bitmaps. Thus, three types of source associated with the employee images were used: photo and two bitmaps that characterize the time of entry and exit from the enterprise.

In the training sample there were 1000 employees, and in the testing – 100 employees. Classes were characterized by the number of days when employees were absent from the workplace (valid and invalid reasons were not differentiated). The division into classes is presented in table 1.
Figure 4. The architecture of the applied CNN.

Table 1. Class distributions

| Class number | 1  | 2    | 3    | 4    | 5          |
|--------------|----|------|------|------|------------|
| Number of missed working days in the next quarter | 0  | 1–2  | 3–5  | 6–8  | More than 8 |

To justify the feasibility of applying the multiple associated images training was carried out for three different situations – when available, all three parameters: X1 – a bit-map input, X2 – a bit-map output, the X3 is the photo. The results of the classifier on the testing sample for three different situations are presented in table 2.

The data in table 2 shows the total number of employees assigned to a particular class and can be used to statistically estimate the loss of working time across the enterprise. However, they do not
provide an indication of the accuracy and completeness of the classification, which can be estimated using the confusion matrix.

Confusion matrix is a $k \times k$ is a size matrix with columns representing actual data and rows representing the results of the classifier.

### Table 2. The results of the classifier

| Class number | The result of the classifier's work for a given composition of associated images (the number of employees assigned to this class) | Actual number of employees assigned to this class |
|--------------|----------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|
|              | X1        | X1, X2 | X1, X2, X3                      |                                                 |
| 1            | 41        | 51     | 74                              | 70                                              |
| 2            | 27        | 21     | 10                              | 14                                              |
| 3            | 21        | 18     | 11                              | 8                                               |
| 4            | 11        | 2      | 4                               | 6                                               |
| 5            | 0         | 8      | 1                               | 2                                               |

When the matrix is filled, the number at the intersection of the row of the class that returned the classifier and the column of the class to which the object really belongs is incremented. In this case, we obtain three matrices for the three compositions of the associated images presented in table 2:

\[
\begin{align*}
CM1 &= \begin{pmatrix}
32 & 8 & 0 & 1 & 0 \\
8 & 12 & 4 & 3 & 0 \\
1 & 7 & 7 & 4 & 2 \\
0 & 2 & 3 & 5 & 1 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}, \\
CM2 &= \begin{pmatrix}
43 & 4 & 2 & 1 & 1 \\
2 & 14 & 3 & 2 & 0 \\
0 & 4 & 8 & 6 & 0 \\
0 & 1 & 0 & 1 & 0 \\
0 & 1 & 2 & 3 & 2
\end{pmatrix}, \\
CM3 &= \begin{pmatrix}
68 & 6 & 0 & 0 & 0 \\
1 & 9 & 0 & 0 & 0 \\
1 & 2 & 8 & 0 & 0 \\
0 & 0 & 0 & 4 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\]

Matrix $CM1$ characterizes the result of classification when training one beat-card $X1$, $CM2$ – bit-cards $X1$ and $X2$ and $CM3$ – $X1$, $X2$ and $X3$. Analysis of the matrix structure shows that with the increase in the number of associated images used for classification, their appearance approaches the diagonal, which may indicate an increase in the quality of classification.

Figure 5 shows the results of the classification. For clarity, the actual number of employees belonging to the corresponding class is connected by a dashed line. Asterisks indicate the number assigned to the class of one parameter $X1$, crosses – $X1$ and $X2$, circles – $X1$, $X2$ and $X3$.

![Figure 5. The Results of the classification](image-url)
Analysis of figure 5 shows that the classification of all three images gives the closest grouping around the dotted line of the actual classification and indicates the feasibility of increasing the groups of associated images when performing classifications.

The number of classes determines the number of outputs of each CNN (in this case, the number 5 in the last block in figure 4), so only this parameter needs to be changed when measuring the number of classes in the network architecture.

5. Conclusion
As a result of this work, the analysis of the current state of application of digital economy technologies in chemical enterprises has been carried out.

Promising directions of digital transformation of the chemical industry and the relevance of the use of deep neural networks in this process are shown.

The scientific novelty of the work consists in the proposed architecture of fuzzy classification system based on the recognition of images associated with the object by convolutional neural networks.

The approbation of the algorithm implementing the proposed fuzzy classifier for the task of estimating the loss of working time at the enterprise during the quarter is carried out. It was found that with an increase in the number of groups associated with the employee images there is a tendency to reduce the classification error. It was found that with an increase in the number of groups associated with the employee images there is a tendency to reduce the classification error.

The proposed algorithm of fuzzy classifier based on networks can be used in decision support systems in various subject areas, where data on the object of study can be presented not only in numerical form, but also in the form of various images.

Acknowledgments
The reported study was funded by RFBR according to the research project № 19-01-00425.

References
[1] Bobkov V I, Borisov V V, Dli M I and Meshalkin V P 2018 Multicomponent Fuzzy Model for Evaluating the Energy Efficiency of Chemical and Power Engineering Processes of Drying of the Multilayer Mass of Phosphorite Pellets Theoretical Foundations of Chemical Engineering. 52 5 pp 786 – 99
[2] Dutta D and Murthy A S 2010 Fuzzy transportation problem with additional restrictions ARPN Journal of Engineering and Applied Sciences 5 2 pp 36 – 40
[3] Lee H, Grosse R, Ranganath R and Ng A Y 2009 Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations Proceedings of the 26th Annual International Conference on Machine Learning pp 609 – 16
[4] Ciresan D M and Schmidhuber J 2012 Multi-column deep neural networks for image classification Preprint arXiv: 1202.2745
[5] Smith S 2008 Digital signal processing: practical handbook for engineers and scientists (Moscow: Dodeka–XXI)
[6] Kotomin A 2012 Recognition of voice command with the use of convolutional neuron networks Digest of annual youth conference: Knowledge–intensive information technologies SIT-2012 (Pereslavl–Zalesky) p. 17–28.
[7] LeCun Y, Kavukcuoglu K and Farabet C 2010 Convolutional networks and applications in vision In Circuits and Systems (ISCAS) Proceedings of 2010 IEEE International Symposium p 253–256
[8] Oppenheim A V and Schafer R W 2004 From frequency to quefrency: a history of the cepstrum IEEE Signal Processing Magazine 21 5 p 95–106
[9] Petrov D A 1986 Two-components and three-components systems (Moscow: Metallurgiya) p 256
[10] Geron Au 2018: The applied machine learning with the help of Scikit-Learn and conceptions, tools and techniques for intetellectual systems creation (Moscow: Dialektika)
[11] Izotov P Suhkonov S and Golovashkin D 2010 The technology of realization of neural networks algorithm in CUDA environment on the example of handwritten digit recognition *Computer optics* **34** 2 pp 243–51.

[12] Candrot E and Sanders J 2016 *CUDA technology on examples Introduction to programming of graphic processes* (Moscow: DMK –Press) p 234