Volodymyr Pustovarov

Kharkiv Representative Office of the General Customer – State Space Agency of Ukraine, Kharkiv, Ukraine

CONSTRUCTION OF INFORMATION TECHNOLOGY FOR DEVELOPMENT OF KNOWLEDGE BASE ON IDENTIFICATION OF URBAN STRUCTURES ON DIGITAL SPACE AND AERIAL PHOTOGRAPHS IN THE URBAN ENVIRONMENT MONITORING

Abstract. The efficiency and quality of modern city management are directly linked to the relevant monitoring. Improving the efficiency and validity of the obtained data on the urban environment is possible due to the automation of the system of urban objects identification on digital space and aerial photographs, which involves determining their changes. Therefore, a knowledge base (data), which consists of a set of rules, facts, an inference mechanism that can be implemented through the use of deep neural network or hybrid (fuzzy neural network) models, is therefore needed. This academic paper proposes the technology of formal presentation of the information technology construction for the development of a knowledge base on the identification of urban structures on digital space and aerial photographs in automated monitoring of the urban environment. The developed technology identifies an interdependent set of phases, with the possibility of further parallelization of sub-stages, taking into account the peculiarities of knowledge representation (formalization) on the identification of urban structures on digital space and aerial photographs in automated monitoring of the urban environment.

Keywords: information technology; knowledge base; deep neural network; fuzzy set; automated monitoring; urban environment.

Introduction

Formulation of the problem. It is known that the implementation of effective management of the modern city requires timely receipt of data, which is ensured by appropriate monitoring [1-5]. One of the main requirements for such monitoring is the efficiency and reliability of the data obtained.

Improving the efficiency and reliability of the data obtained in monitoring the urban environment is closely related to the development of an automated system used for identifying urban objects in digital space and aerial photographs, determining the presence of their changes as an intelligent automated system.

The main element of automated systems for identifying urban objects can be a knowledge base (KB), which is a set of rules, facts, inference mechanisms, implemented through the use of a deep neural network or hybrid (fuzzy neural network) models.

Development of a knowledge base is one of the most time-consuming phases in the creation of intelligent systems. However, the issues of an integrated approach to the development of a knowledge base at the industrial level to present and accumulate knowledge on the processes of identifying urban structures on digital space and aerial photographs using deep neural networks and hybrid (fuzzy neural networks) models are insufficiently studied.

Thus, in the subject area there is a contradiction, which, on the one hand, consists in the need to develop a knowledge base (data) for automated systems of identifying urban structures on digital space and aerial photographs using deep neural networks and fuzzy logic, on the other hand the above contradiction consists in the limited capabilities of existing technologies used for the development of such a class of KB with a comprehensive presentation and acquisition of knowledge on the processes of identifying certain objects on digital space and aerial photographs, and the processes of detecting changes in the urban environment as well as carrying out analysis of their causes.

Analysis of recent research and publications. Currently, there is a substantial body of scholarly work on the development of KB for various purposes [6-8]. Traditionally, the basic technology of KB creation is based on classical technologies of intelligent systems development.

In this case, the prototyping method, which is based on the implementation of a certain sequence of phases of KB development with an intermediate presentation, refinement and creation of a relevant prototype, is used as a basic approach [9]. However, the phases of classical technologies for the creation of intelligent systems are characterized by a fairly conditional, non-formalized, definition of the boundaries of the above stages development, as well as the transitions between them.

All this brings much complexity to the technology of creating a prototype KB for monitoring the urban environment from an industrial perspective, does not make it possible to minimize the required amount of work on the creation of a knowledge base, poses great challenges to the development management, complicates the formulation of tasks and functional responsibilities of developers.

The purpose of the article is a develop a formalized presentation of the construction of information technology for the development of KB on the identification of urban structures on digital space and aerial photographs in automated monitoring of the urban environment using deep neural networks and fuzzy logic.

Basic material

Formally, the information technology of KB development on the identification of urban structures on digital space and aerial photographs in automated

monitoring of the urban environment using deep neural networks and fuzzy logic is proposed to be presented in the form of a tuple [10-12]:

$$T^{KBF} = \left\langle P^{KBF}, \left\{ S_i^{KBF} \right\}, \left\{ C_j^{KBF} \right\}, \left\{ L_k^{KBF} \right\}, \left\{ M_n^{KBF} \right\}, \right\rangle \quad (1)$$

where P^{KBF} is setting the task of KB development on the identification of urban structures on digital space and aerial photographs;

$\left\{ S_i^{KBF} \right\}$ is a set of stages of the corresponding KB development, $i = \overline{1, I}$, I is the total number of development stages;

$\left\{ C_j^{KBF} \right\}$ is a set of connections between the stages from the set $\left\{ S_i^{KBF} \right\}$, $j = \overline{1, J}$, J is the total number of connections;

$\left\{ L_k^{KBF} \right\}$ is a plethora of languages for representation of knowledge on the identification of urban structures on digital space and aerial photographs, $k = \overline{1, K}$, K is the total plethora of knowledge representation languages;

$\left\{ M_n^{KBF} \right\}$ is a set of models of knowledge representation, $n = \overline{1, N}$, N is the total number of models.

The modified phases of classical intelligent systems development are considered as the basic phases from the set $\left\{ S_i^{KBF} \right\}$:

– the stage of knowledge identification $S_1^{KBF} \in \left\{ S_i^{KBF} \right\}$ on the identification of urban structures on digital space and aerial photographs;

– the phases of knowledge conceptualization $S_2^{KBF} \in \left\{ S_i^{KBF} \right\}$;

– the phases of knowledge formalization $S_3^{KBF} \in \left\{ S_i^{KBF} \right\}$;

– the phases of software implementation (prototyping) $S_4^{KBF} \in \left\{ S_i^{KBF} \right\}$;

– the phases of software verification of KB $S_5^{KBF} \in \left\{ S_i^{KBF} \right\}$ [13].

On the phase of knowledge identification S_1^{KBF}, the general statement of the problem of knowledge base development, which determines the necessary resources and sources of knowledge, is achieved.

On the phase of conceptualization S_2^{KBF}, the conceptual model (field of knowledge) of KB $M_1^{KBF} \in \left\{ M_n^{KBF} \right\}$ is formed subject to the following sub-stages:

1) based on the analysis of the subject area, the composition of basic knowledge is determined in the form of a verbal description (composition of the field of knowledge) on the identification of urban structures on digital space and aerial photographs in automated monitoring of the urban environment – $S_2^{K(KBF)}$;

2) based on the knowledge composition, the field of knowledge (structure of the subject area) is directly formed using the appropriate language of the field of knowledge description (this technology uses the construction of conceptual models based on the language of object-oriented modeling UML $L_m^{KBF} \in \left\{ L_m^{KBF} \right\}$ [14]) – $S_2^{M_i(KBF)}$.

On the phase of (presentation) knowledge formalization S_3^{KBF} – a mathematical $M_2^{KBF} \in \left\{ M_n^{KBF} \right\}$ and logical $M_3^{KBF} \in \left\{ M_n^{KBF} \right\}$ model of KB is formed subject to the following sub-stages [15-18]:

1) the formalized composition of knowledge on the identification of urban structures on digital space and aerial photographs and the presence of changes in the urban environment are determined – $S_3^{KS(KBF)}$;

2) methods of knowledge representation and the choice (development) of methods of knowledge formalization are determined – $S_3^{KR(KBF)}$. In particular, to present knowledge on the identification of urban structures on digital space and aerial photographs and the presence of changes in the urban environment, it is proposed to apply the methods based on the use of interval fuzzy sets of the second type $L_2^{KBF} \in \left\{ L_k^{KBF} \right\}$, interval fuzzy logic systems of the second type $L_3^{KBF} \in \left\{ L_k^{KBF} \right\}$ and modified deep neural networks (convolutional neural network) based on pre-trained models, as well as the models for semantic segmentation $L_3^{KBF} \in \left\{ L_k^{KBF} \right\}$;

3) a directly formalized description of knowledge (formation of the mathematical model M_2^{KBF}) is performed using the selected methods of presentation as well as the developed methods of knowledge formalization – $S_3^{KD(KBF)}$;

4) for each formalized subtask (segmentation of urban structures on digital space and aerial photographs, detection of changes in the urban environment) a generalized algorithm for its solution is developed (within the mathematical model M_2^{KBF}), the formalized knowledge is taken into account and the trends in the process of KB construction and presentation of the results of solving subtasks are detected – $S_3^{AM(KBF)}$;

5) the elements of the mathematical model , M_2^{KBF} which are considered as procedural knowledge, are transformed into a logical model of the KB M_3^{KBF}, described using the constructions of representation of

logical models of the language of object-oriented modeling UML $L_1^{KBF} - S_3^{LM(KBF)}$.

On the phase of software implementation (preliminary prototyping) of the KB S_4^{KBF}, a physical model of the KB $M_4^{KBF} \in \left\{ M_n^{KBF} \right\}$ is constructed , and thus, the software prototype of the KB when performing the following sub-stages is likewise created:

1) the KB architecture, which determines the structure, functions and relationships of the components of the KB – $S_4^{BA(KBF)}$ is developed. In this case, the main elements of the KB are the membership functions of fuzzy sets, the training sample (implicit KB) and the synaptic maps of neural networks;

2) the implementation (prototyping) of the components of the KB – $S_4^{BP(KBF)}$ takes place. In this case, the implementation of the KB components includes the development of design options (model formation M_4^{KBF}) for software implementation of the KB using physical language models of the object-oriented modeling of UML and direct software implementation of the KB using a high-level programming language $L_6^{KBF} \in \left\{ L_k^{KBF} \right\}$ and (or) the KB (data) management system $L_7^{KBF} \in \left\{ L_k^{KBF} \right\}$;

3) the process of populating the KB is based on the elements of the mathematical model M_2^{KBF} , which are considered as declarative knowledge – $S_4^{BP(KBF)}$. The populating of the KB is carried out both as a result of direct introduction of declarative knowledge into the database (for example, when introducing a new or expanding the existing training sample) and as a result of learning the elements of the KB (for example, synaptic maps are generated through neural networks training).

Knowledge can be represented in the form of synaptic maps, which are proposed to be determined by the results of construction and further training of neural networks using specialized simulation tools.

Testing (verification) of the KB $S_5^{KBF} \in \left\{ S_i^{KBF} \right\}$ is performed directly either on the design solutions,

presented in the form of UML models, or on the program code that implements the prototype (or final version) of the KB on recognition.

In case of errors, the return can be made to any previous stage of KB development, depending on the type of error detected.

Therefore, the set of connections $\left\{ C_j^{KBF} \right\}$ between the stages of the set $\left\{ S_i^{KBF} \right\}$ is determined by the order of the above stages implementation (direct connections $C_1^{KBF} - C_4^{KBF}$) and the test results (feedback) $C_5^{KBF} - C_8^{KBF}$.

Thus, multiple phases of information technology T^{KBF} is set as follows:

$$\left\{ S_i^{KBF} \right\} = \left\{ S_1^{KBF}, S_2^{KBF}, S_3^{KBF}, S_4^{KBF}, S_5^{KBF} \right\}, \quad (2)$$

$$S_2^{KBF} = S_2^{K(KBF)} \bigcup S_2^{M(KBF)} , \quad (3)$$

$$S_3^{KBF} = S_3^{KS(KBF)} \bigcup S_3^{KR(KBF)} \bigcup S_3^{KD(KBF)} \bigcup$$
$$\bigcup S_3^{AM(KBF)} \bigcup S_3^{LM(KBF)}; \quad (4)$$

$$S_4^{KBF} = S_4^{BA(KBF)} \bigcup S_3^{BP(KBF)} \bigcup S_3^{BK(KBF)} . \quad (5)$$

Conclusions

Thus, the construction of information technology for the development of KB with regard to the identification of urban structures on digital space and aerial photographs in automated monitoring of the urban environment is proposed. Information technology makes it possible to implement the industrial development of KB on the identification of urban structures on digital space and aerial photographs using neural networks and fuzzy logic, with a clearly defined and interdependent set of phases. Whereby, it is possible to further parallelize the implementation of sub-phases, taking into account the peculiarities of knowledge representation (formalization) on the identification of urban structures on digital space and aerial photographs in automated monitoring of the urban environment. Prospects for further research in this area may be the development of proposals for the use of CAD-facilities to design KB in automated systems for urban environment monitoring.

REFERENCES

1. Shikhov, A.N., Cherepanova, E.S. and Ponomarchuk, A.I. (2014), *Geoinformation systems: the use of GIS technologies in solving hydrological problems, a workshop, textbook. manual*, Perm, 91 p.
2. De Mers and Michael, N. (1999), *Geographic information systems. Foundations*, Date+, Moscow, 506 p.
3. Andrieiev, S. and Zhilin, V. (2020), "Methodology of creation of atlases of historical cartographic models from data of aerophotography with the use of геоінформаційних technologies", *Advanced Information Systems*, Vol. 4, No. 1, pp. 45-62, DOI: http://dx.doi.org/10.20998/2522-9052.2020.1.08
4. Andrieiev, S., Zhilin, V. and Melnyk, A. (2019), "The use of anamorphosis cartographic models for geodata analysis", *Advanced Information Systems*, Vol. 3, No. 3, pp. 5-16, DOI: http://dx.doi.org/10.20998/2522-9052.2019.3.01
5. Andreev, S. and Zhilin, V. (2019), "Application of aerophotic data with unmanned aircraft for developing 3D models of terrain", *Control, navigation and communication systems*, No. 1(53), Poltava NTU Yuri Kondratyuk, Poltava, pp. 3-16, DOI: http://dx.doi.org/10.26906/SUNZ.2019.1.003
6. Dennis, A., Wixom, B. & Tegarden, D. (2015), *Systems analysis and design: An object-oriented approach with UML*, John Wiley & Sons, NY, 525 p.

7. Bonczek, R.H., Holsapple, C. and Whinston, A.B. (1981), *Foundations of Decision Support Systems*, Acad. Press, New York.

8. Anisimova, O.L., Zraenko, D.Yu. and Komosko, V.V. (2010), "Storage of spatial objects as a part of the regional node of IPD UrFO: model of storage", *Spatial data*, No. 1, p. 62.

9. Inmon, W.H. (1991), *Building the Data Warehouse*, John Wiley & Sons, New-York, 312 p.

10. Weizmann, V.M. (2002), *Design of economic information systems*, MUBiNT, Yaroslavl, 213 p.

11. Terelyansky, P.V. (2009), *Decision support systems. Experience of design*, Volga State Techn. University, Volgograd, 127 p.

12. Shannon, K. (1963), *Works on information theory and cybernetics*, IIL, Moscow, 829 p.

13. Spirley, E. (2001), *Corporate data warehouses. Planning, development, implementation*, Williams, Moscow, 400 p.

14. Ruban, I., Kuchuk, H. and Kovalenko A. (2017), "Redistribution of base stations load in mobile communication networks", *Innovative technologies and scientific solutions for industries*, No. 1 (1), pp. 75-81, DOI : https://doi.org/10.30837/2522-9818.2017.1.075.

15. Kovalenko, A. and Kuchuk H. (2018), "Methods for synthesis of informational and technical structures of critical application object's control system", *Advanced Information Systems*, Vol. 2, No. 1, pp. 22–27, DOI: https://doi.org/10.20998/2522-9052.2018.1.04.

16. Nechausov A., Mamusûc I., Kuchuk N. (2017), "Synthesis of the air pollution level control system on the basis of hyperconvergent infrastructures", *Advanced Information Systems*, Vol. 1, No. 2. P. 21 – 26. DOI: https://doi.org/10.20998/2522-9052.2017.2.04.

17. Yaloveha, V., Hlavcheva, D. and Podoroxhniak, A. (2019), "Usage of convolutional neural network for multispectral image processing applied to the problem of detecting fire hazardous forest areas", *Advanced Information Systems*, Vol. 3, No. 1, pp. 116–120, DOI: https://doi.org/10.20998/2522-9052.2019.1.19

18. Ronneberger, O., Fischer, P. & Brox, T. (2015), "U-net: Convolutional networks for biomedical image segmentation", *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp. 234-241.

Відомості про авторів / About the Authors

Пустоваров Володимир Володимирович – начальник групи харківського представництва генерального замовника – Державного космічного агентства України, Харків, Україна;

Volodymyr Pustovarov – chief of group of the Kharkov representative office of general customer – the State space agency of Ukraine, Kharkiv, Ukraine;

e-mail: pustovarov78volodymyr@gmail.com; ORCID ID: https://orcid.org/0000-0003-3944-5771.

Побудова інформаційної технології розробки бази знань про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища

В. В. Пустоваров

А н о т а ц і я . Ефективність і якість управління сучасним містом безпосередньо пов'язані з веденням відповідного моніторингу. Підвищення оперативності і достовірності отриманої інформації про міське середовище можливе завдяки автоматизації системи розпізнавання міських об'єктів на цифрових космічних та аерофотознімках з визначенням їх змін. Отже, необхідна база знань (даних), яка складена із сукупності правил, фактів, механізмів виведення, що можуть бути реалізовані на основі використання глибоких нейромережевих або гібридних (нечітких нейромережевих) моделей. В статті запропоновано технологію формального представлення побудови інформаційної технології розробки бази знань про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища. Розроблена технологія визначає взаємозалежну сукупність етапів, з можливістю подальшого розпаралелювання виконання підетапів, з урахуванням особливостей подання (формалізації) знань про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища.

К л ю ч о в і   с л о в а : інформаційна технологія; база знань; глибока нейронна мережа; нечітка множина; автоматизований моніторинг; міське середовище.

Построение информационной технологии разработки базы знаний о распознавании городских строений на цифровых космических и аэрофотоснимках при автоматизированном мониторинге городской среды

В. В. Пустоваров

А н н о т а ц и я . Эффективность и качество управления современным городом непосредственно связанны с ведением соответствующего мониторинга. Повышение оперативности и достоверности полученной информации о городской среде возможно благодаря автоматизации системы распознавания городских объектов на цифровых космических и аэрофотоснимках с определением их изменений. Следовательно, необходима база знаний (данных), которая составлена из совокупности правил, фактов, механизмов вывода, которые могут быть реализованы на основе использования глубоких нейросетевых или гибридных (нечетких нейросетевых) моделей. В статье предложена технология формального представления построения информационной технологии разработки базы знаний о распознавании городских строений на цифровых космических и аэрофотоснимках при автоматизированном мониторинге городской среды. Разработанная технология определяет взаимосвязанную совокупность этапов, с возможностью дальнейшего распараллеливания выполнения подэтапов, с учетом особенностей представления (формализации) знаний о распознавании городских строений на цифровых космических и аэрофотоснимках при автоматизированном мониторинге городской среды.

К л ю ч е в ы е   с л о в а : информационная технология; база знаний; глубокая нейронная сеть; нечеткое множество; автоматизированный мониторинг; городская среда.