

Труды МАИ. 2023. № 130

Trudy MAI, 2023, no. 130

## **ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ**

Научная статья

УДК 004.01

DOI: [10.34759/trd-2023-130-18](https://doi.org/10.34759/trd-2023-130-18)

### **АЛГОРИТМ СЕМАНТИЧЕСКОЙ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПОЗИЦИОНИРОВАНИЯ ЛЕТАТЕЛЬНОГО АППАРАТА НА ЗЕМНОЙ ПОВЕРХНОСТИ**

**Дарья Сергеевна Олькина**

Московский физико-технический институт (национальный исследовательский университет), МФТИ,

Долгопрудный, Московская область, Россия

[olkina.ds@phystech.edu](mailto:olkina.ds@phystech.edu)

***Аннотация.*** В данной работе рассматриваются методы определения подхода к разработке алгоритма сегментации, а также выбор отличительных признаков для решения задачи сегментации изображений земной поверхности для определения координат беспилотного летательного аппарата (БЛА) с целью его позиционирования на земной поверхности. Теоретической значимостью исследования является численное исследование постановок задач сегментации изображений стохастическими методами. Практической значимостью является

получение результатов экспериментов по сегментации изображений видимого и радио-диапазонов, разработка инструментов обучения и программного комплекса, решающего задачи сегментации многоканальных изображений.

**Ключевые слова:** компьютерное зрение, сверточная нейронная сеть, сегментация, архитектура Visenet, режим реального времени, аналог инерциальной навигации.

**Благодарности:** Автор выражает благодарность доценту кафедры прикладной механики и информатики, кандидату физико-математических наук, Дорофееву Евгению Александровичу за постановку задачи и интерес к работе.

**Для цитирования:** Олькина Д.С. Алгоритм семантической сегментации изображений для решения задачи позиционирования летательного аппарата на земной поверхности // Труды МАИ. 2023. № 130. DOI: [10.34759/trd-2023-130-18](https://doi.org/10.34759/trd-2023-130-18)

## **INFORMATICS, COMPUTATION ENGINEERING AND MANAGEMENT**

Original article

### **ALGORITHM OF SEMANTIC IMAGE SEGMENTATION FOR SOLVING THE PROBLEM OF POSITIONONG AN AIRCRAFT ON THE EARTH'S SURFACE**

**Daria S. Olkina**

Moscow Institute of Physics and Technology,

Dolgoprudny, Moscow Region, Russia

[olkina.ds@phystech.edu](mailto:olkina.ds@phystech.edu)

**Abstract.** This work is devoted to one of the main methodological issues of using the object-oriented approach – choosing the method of segmentation of multichannel images. The methods for determining the approach to the development of a segmentation algorithm are considered, as well as the choice of distinctive features to solve the problem of segmentation of images of the earth's surface to determine the coordinates of an unmanned aerial vehicle (UAV) in order to position it on the earth's surface.

The theoretical significance of the study is a numerical study of the formulation of the tasks of segmentation of images by stochastic methods. Of practical importance is obtaining the results of experiments on the segmentation of images of the visible and radio ranges, development of training tools and a software package that solves the problems of segmentation of multichannel images.

The main objectives of the study are to develop a model of neural network segmentation of multichannel images of optical and radio range, construction of the procedure for training segmentation models, obtaining comparative estimates of the computational complexity of algorithms and learning characteristics, hyperparametric optimization of various neural network models, obtaining the main dependencies of training parameters, quality and speed of models. The work uses the neural network of semantic segmentation Bisenet, which allows you to get segmented images of the earth's surface in real time. For the neural network Bisenet, some hyperparameters were configured: the most suitable optimizer and the best strategy for reducing the learning speed were identified (learning rate).

**Keywords:** computer vision, convolutional neural network, segmentation, Bisenet architecture, real-time mode, analogue of inertial navigation

**Acknowledgements:** The author expresses his gratitude to the associate professor of the Department of Applied Mechanics and Informatics of MIPT, Candidate of Physical and Mathematical Sciences Evgeny Alexandrovich Dorofeev for the formulation of the problem and interest in the work.

**For citation:** Olkina D.S. Algorithm of semantic image segmentation for solving the problem of positioning an aircraft on the earth's surface. *Trudy MAI*, 2023, no. 130. DOI: [10.34759/trd-2023-130-18](https://doi.org/10.34759/trd-2023-130-18)

## Введение

Впервые вопрос развития компьютерного зрения начали изучать в 1950 году. Алан Тьюринг в своей работе «Вычислительные машины и разум» [3] описал подход, суть которого заключалась в изучении вопроса имитации поведения человека искусственным интеллектом. По причине того, что на тот момент еще не было разработано настоящих компьютеров имитация поведения человека обсуждалась как абстракция на основе сенсорных данных и носила сугубо теоретический характер. Понятие искусственного интеллекта ввел Джон Маккарти в 1955 году [4]. Искусственный интеллект позволяет выявлять сложные зависимости между входными данными и результатом, и анализировать полученную информацию и использовать выводы для достижения оптимальных решений.

В настоящее время с развитием информационных технологий широко применяются методы компьютерного зрения в различных задачах обработки изображений. Компьютерное зрение позволяет решать такие задачи, как

распознавание объектов, сегментация изображений, классификация и кластеризация множества объектов.

В работе Л.А. Демидовой, Н.И. Нестерова, Р.В. Тишкина [1] задача обработки изображений земной поверхности с применением искусственных нейронных сетей решается с помощью разработки новой интеллектуальной технологии сегментации гиперспектральных изображений, основанной на комплексном использовании алгоритмов кластеризации. Для сегментации изображения используются алгоритмы FCM, PCM, PFCM. FCM-алгоритм не всегда точно оценивает координаты центров множества объектов, PCM-алгоритм может приводить к формированию совпадающих множеств объектов [5, 6]. Учитывая недостатки алгоритмов FCM, PCM в работе используется PFCM-алгоритм. Для решения нашей задачи данный алгоритм не подходит, так как для сегментации используется только пространственная информация.

В работе С.В.Шелега для решения задачи сегментации изображений в качестве основы алгоритма распознавания используется сверточная нейронная сеть с архитектурой U-Net [2]. U-Net – нейронная сеть семантической сегментации, для обучения которой используется матрица весов размера  $n \times n$ , которую перемещают по всему обрабатываемому слою (на входе сети - непосредственно по входному изображению). Слой свертки производит суммирование результатов поэлементного произведения каждого фрагмента изображения на матрицу (ядро свертки). Весовые коэффициенты ядра свертки неизвестны и устанавливаются в процессе обучения [7]. В отличие от нашего подхода данная нейронная сеть не работает в режиме реального времени.

Как показано ранее, разработка алгоритма семантической сегментации с использованием искусственной нейронной сети для обработки изображений земной поверхности с условием работы алгоритма в режиме реального времени является актуальной задачей.

В данной работе предлагается подход к разработке алгоритма сегментации, включающего в себя методы кластеризации и детекции, и выбора отличительных признаков для решения задачи сегментации изображений земной поверхности для позиционирования летательного аппарата на земной поверхности.

### **Алгоритм сегментации**

При выборе алгоритма сегментации основными задачами являются: выбор типа нейронной сети, алгоритма обучения и формирование обучающего множества. Обучающее множество или выборка состоит из изображений земной поверхности, снятых с беспилотного летательного аппарата и со снимков спутниковых карт с открытых источников. В качестве классов были выбраны: лес, поле, дорога, город, вода – объекты, практически исчерпывающе, описывающие местность в средней полосе. Разметка изображений выполняется в программе для ручного выделения классов на изображении CVAT.

Для обработки изображений земной поверхности выбор типа нейронной сети основывался на необходимости обработки изображений в режиме реального времени, и при этом, предоставляющая высокую точность и качество сегментации, поскольку от последнего зависит точность определения координат. Под данные критерии выбрана нейронная сеть семантической сегментации Bisenet [8].

Нейронная сеть Visenet в отличие от классического подхода семантической сегментации при обработке изображения не выполняет его кодирование с увеличением числа каналов и декодирование, а использует не только смысловую (наименование объектов на изображении) информацию, которую мы получаем множеством сверток при кодировании, а также пространственную информацию (местоположение объектов на изображении и форма объектов).

Входное изображение поступает в 2 пути: пространственный и смысловой (контекстный). Пространственный путь состоит из 3х сверточных блоков, которые включает в себя свертки ядром 3x3 со сдвигом на 2 шага, пакетной нормализацией и нелинейностью Relu. Изображение сжимается 8 раз для получения набора матриц, в которых сохраняется пространственная информация об изображении.

Смысловой путь состоит из нескольких сильно сжимающихся сверточных блоков, уточняющего модуля ARM, который преобразует набор матриц в вектор при помощи усреднения матриц. Полученный вектор искусственной нейронной сети проходит через один полносвязный слой, перемножается со входом ARM и на выходе получаем смысловую информацию об объектах. Выходы контекстного и пространственного путей переходят в модуль слияния признаков.

### **Обучение нейронной сети**

Для обучения нейронной сети Visenet за стартовые веса взяты веса, предобученные на наборе данных ImageNet и выполнялась настройка гиперпараметров сети: был выявлен наиболее подходящий оптимизатор, лучшая стратегия оптимизации темпа обучения и подбор коэффициента скорости обучения. Схема построения нейронной сети Visenet представлена на Рис.1.

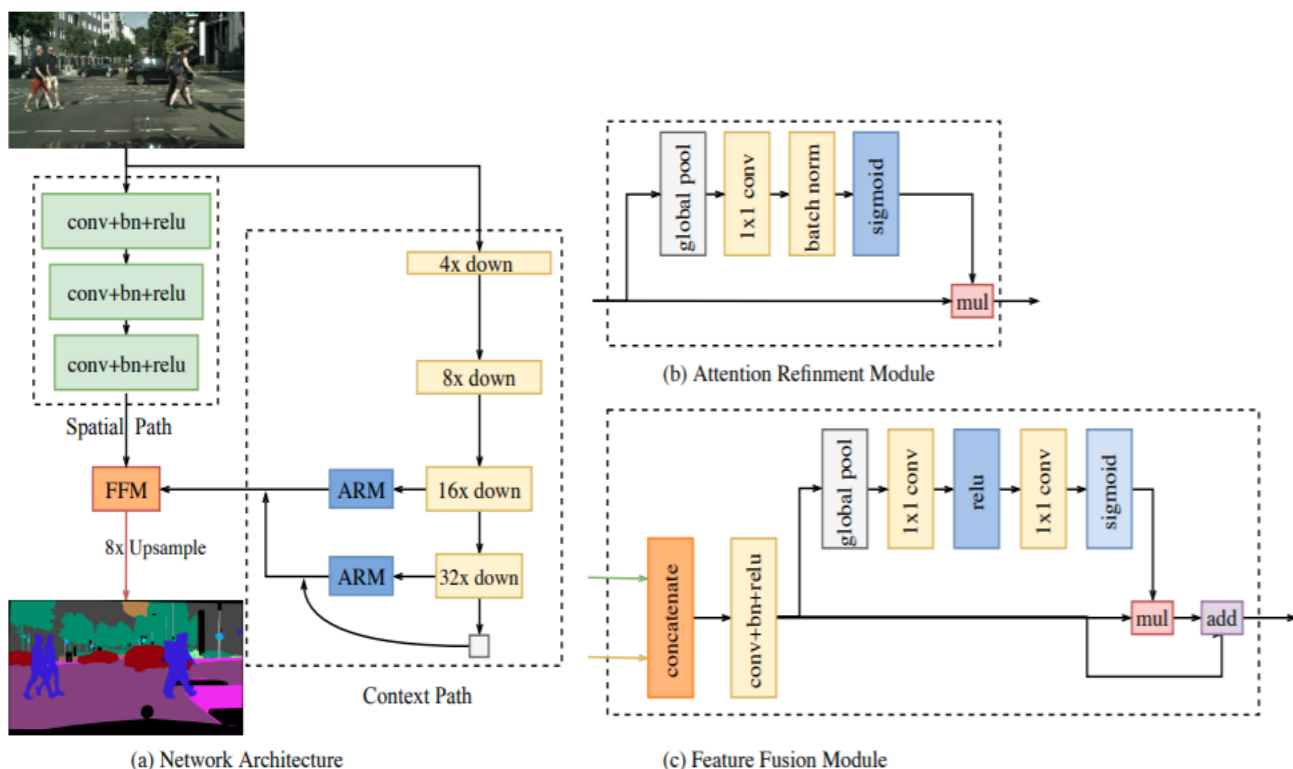


Рис. 1. Схема построения нейронной сети Bisenet

### Решение задачи позиционирования летательного аппарата

Для решения задачи позиционирования летательного аппарата выполнялась склейка изображений для получения панорамы. Для этого на снимках выделялись особые точки методом ORB [9], затем наборы точек сопоставлялись методом перебора и выполнялся поиск изображения методом гомографии [10].

Поиск сегментированной панорамы на предварительно сегментированной карте местности происходил с помощью библиотечных алгоритмов opencv методом template matching (как и при свертке изображений сегментированная панорама скользит по карте и вычисляются коэффициенты корреляции, место максимального коэффициента – наиболее вероятное положение панорамы).

Таким образом мы получаем координаты центра панорамы, но учитывая, что БЛА в этот момент находится приблизительно над центром последнего из снимков



этой панорамы, стоит внести к ним поправку, вычисляемую при получении панорамы.

Необходимо также учесть и углы тангажа и крена БЛА. На Рис. 2 представлен пример сопоставления особых точек при сшивании панорамы.

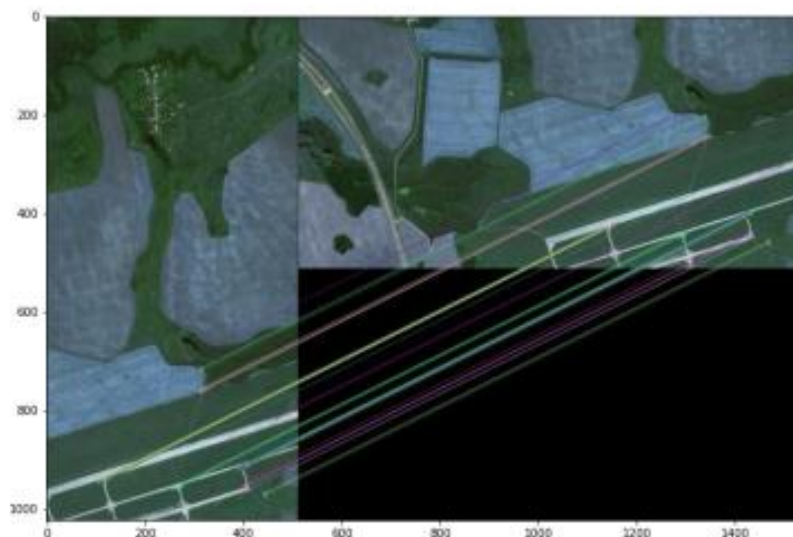


Рис. 2. Пример сопоставления особых точек при сшивании панорамы

### **Заключение**

В данной работе, для решения задачи позиционирования ЛА на местности использован алгоритм семантической сегментации с использованием искусственной нейронной сети V-Net для обработки изображений земной поверхности. Для повышения точности позиционирования ЛА на местности на глобальной карте рассматриваются не только отдельные снимки местности, но и сегментированная панорама, полученная из данных снимков. Пример сегментации снимка представлен на Рис.3:

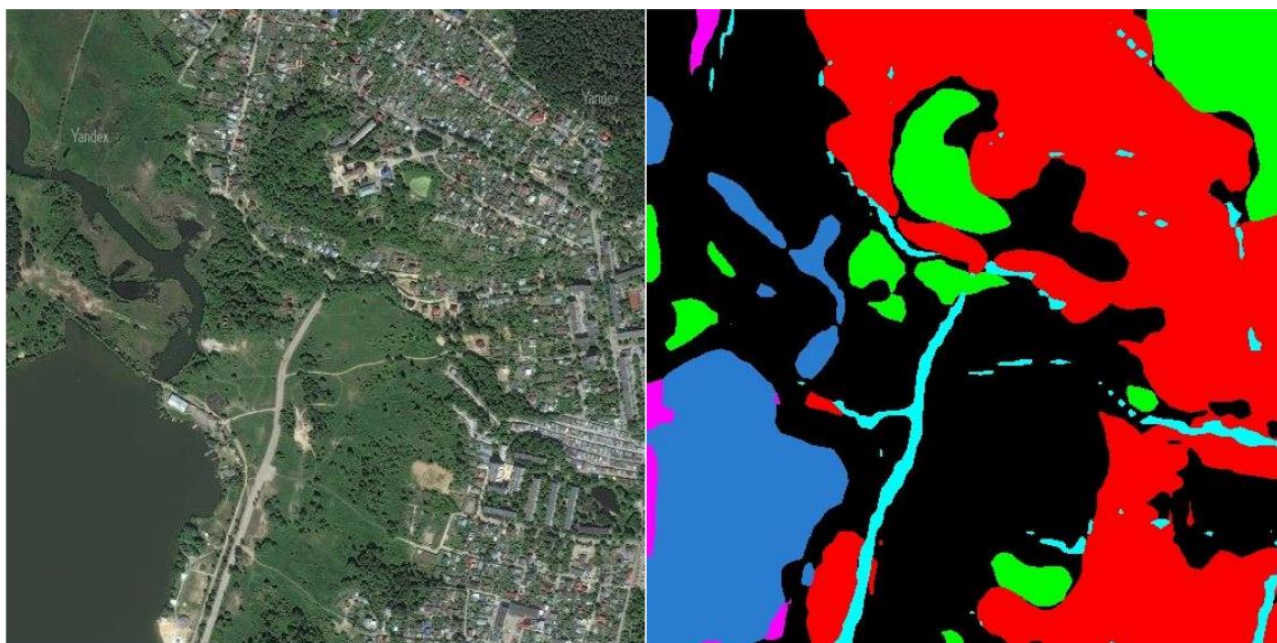


Рис. 3. Пример сегментации снимка

Данный алгоритм может применяться как аналог инерциальной навигации, основанный на подсчете перемещения ЛА на каждом временном шаге при помощи вычисления сдвига между последовательными снимками местности.

### **Список источников**

1. Демидова Л.А., Нестеров Н.И., Тишкин Р.В. Возможность-нечеткая сегментация изображений земной поверхности с применением генетических алгоритмов и искусственных нейронных сетей // Информатика. Телекоммуникации. Управление. 2014. № 3 (198). С. 37-47.
2. Шелег С.В. Семантическая сегментация аэрофотоснимков сельскохозяйственной растительности на базе сверточных нейронных сетей: автореф. дисс. ... магистра технических наук. – Минск: БГУИР, 2021. – 9 с.

3. Алан Тьюринг. Вычислительные машины и разум / пер. с англ. – М.: Изд-во АСТ, 2018. - 128 с.
4. Клиффорд Пиквер. Искусственный интеллект. Иллюстрированная история. От автоматов до нейросетей / пер. с англ. – М.: Изд-во Синдбад, 2021. - 224 с.
5. Демидова Л.А., Кираковский В.В. Кластеризация объектов на основе нечетких множеств второго типа и генетического алгоритма // Труды научно-практической конференции «Управление созданием и развитием систем, сетей и устройств телекоммуникаций (Санкт-Петербург, 2008). - СПб.: НОЦ "Перспектива", 2008.
6. Rhee F. C.-H. Uncertain fuzzy clustering: insights and recommendations // IEEE Computational intelligence magazine, 2007, vol. 2, no. 1, pp. 44-56.  
DOI:[10.1109/MCI.2007.357193](https://doi.org/10.1109/MCI.2007.357193)
7. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. - СПб.: Издательский дом «Питер», 2018. - 482 с.
8. Changqian Yu, Changxin Gao, Jingbo Wang, Gang Yu, Chunhua Shen, Nong Sang. BiSeNet V2: Bilateral Network with Guided Aggregation for Real-time Semantic Segmentation // International Journal of Computer Vision, 2020, vol. 129 (11), pp. 1-18.  
DOI:[10.1007/s11263-021-01515-2](https://doi.org/10.1007/s11263-021-01515-2)
9. Yingxiao Xu, Long Pan, Chun Du, Jun Li, Ning Jing, and Jiangjiang Wu. Vision-based UAVs Aerial Image Localization: A Survey. 2018. DOI: [10.1145/3281548.3281556](https://doi.org/10.1145/3281548.3281556)
10. Ефимов А.И., Новиков А.И. Алгоритм поэтапного уточнения проективного преобразования для совмещения изображений // Компьютерная оптика. 2016. Т. 40. № 2. С. 258–265.

11. Ермаков П.Г., Гоголев А.А. Сравнительный анализ схем комплексирования информации бесплатформенных инерциальных навигационных систем беспилотных летательных аппаратов // Труды МАИ. 2021. № 117. URL: <https://trudymai.ru/published.php?ID=156253>. DOI: 10.34759/trd-2021-117-11
12. Ивашова Н.Д., Михайлин Д.А., Чернякова М.Е., Шаныгин С.В. Нейросетевое решение задачи оперативного планирования маршрутного полета беспилотных летательных аппаратов и назначение времени наблюдения наземных объектов с помощью нечеткой логики при отображении этих результатов на экране компьютера до вылета // Труды МАИ. 2019. № 104. URL: <https://trudymai.ru/published.php?ID=102223>
13. Соколов Д.Ю. Сравнительный анализ метода применения искусственной нейронной сети в целях решения задачи инженерно-штурманского расчета полета летательного аппарата // Вооружение и экономика. 2012. № 1 (17). С. 50-57.
14. Соколов Д.Ю. Применение искусственной нейронной сети для решения задач прогнозирования движения наземных объектов // Труды МАИ. 2022. № 123. URL <https://trudymai.ru/published.php?ID=165563>. DOI: 10.34759/trd-2022-123-17
15. Гусейнов Г.А., Зульфугарлы П.Р., Абдуррахманова И.Г. О точности прямого геореференцирования беспилотного летательного аппарата в зонах с различными климатическими условиями // Труды МАИ. 2022. № 126. URL: <https://trudymai.ru/published.php?ID=169012>. DOI: 10.34759/trd-2022-126-26
16. Чигринец Е.Г., Верченко А.В. CAD/CAM/CAE системы, OMW-технологии и нейросетевые алгоритмы анализа данных на предприятиях авиастроительной

отрасли // Труды МАИ. 2019. № 104. URL:  
<https://trudymai.ru/published.php?ID=102420>

17. Шейников А.А., Коваленко А.М., Санько А.А. Точность определения координат беспилотного летательного аппарата с навигационным комплексом, включающим оптико-электронную систему позиционирования // Научный вестник МГТУ ГА. 2023. Т. 26. № 1. С. 81-94. URL: <https://doi.org/10.26467/2079-0619-202-26-1-81-94>

18. Степанов Д.Н., Тищенко И.П. Задача моделирования полета беспилотного летательного аппарата на основе системы технического зрения // Программные системы: теория и приложения. 2011. № 4. С. 33-43.

19. Степанов Д.Н., Тищенко И.П. Программная система для моделирования полета БПЛА и его автономного позиционирования с использованием методов и алгоритмов технического зрения // XIV национальная конференция по искусственному интеллекту с международным участием, КИИ-2014 (Казань, 24-27 октября 2014). – Казань: Изд-во РИЦ «Школа», 2014. Т. 3. С. 333-341.

20. Конушин А. Геометрические свойства нескольких изображений // Компьютерная графика и мультимедиа. 2006. № 4 (3). URL: <http://archive.ec/hwXHz>

## References

1. Demidova L.A., Nesterov N.I., Tishkin R.V. *Informatika. Telekommunikatsii. Upravlenie*, 2014, no. 3 (198), pp. 37-47.
2. Sheleg S.V. *Semanticheskaya segmentatsiya aerofotosnimkov sel'skokhozyaistvennoi rastitel'nosti na baze svertochnykh neironnykh setei* (Semantic segmentation of aerial

photographs of agricultural vegetation on the basis of convolutional neural networks):  
autoref. diss. ... Master of Technical Sciences. Minsk, BGUIR, 2021, 9 p.

3. Alan T'yuring. *Vychislitel'nye mashiny i razum* (Computers and the Mind), Moscow, Izd-vo AST, 2018, 128 p.

4. Klifford Pikover. *Iskusstvennyi intellekt. Illyustrirovannaya istoriya. Ot avtomatov do neirosetei* (Artificial intelligence. Illustrated story. From automata to neural networks), Moscow, Izd-vo Sindbad, 2021, 224 p.

5. Demidova L.A., Kirakovskii V.V. *Trudy nauchno-prakticheskoi konferentsii «Upravlenie sozdaniem i razvitiem sistem, setei i ustroystv telekommunikatsii, Saint-Petersburg, NOTs "Perspektiva", 2008.*

6. Rhee F. C.-H. Uncertain fuzzy clustering: insights and recommendations, *IEEE Computational intelligence magazine*, 2007, vol. 2, no. 1, pp. 44-56.  
DOI:10.1109/MCI.2007.357193

7. Nikolenko S., Kadurin A., Arkhangel'skaya E. *Glubokoe obuchenie. Pogruzhenie v mir neironnykh setei* (Deep learning. Immersion in the world of neural networks), Saint Petersburg, Izdatel'skii dom «Piter», 2018, 482 p.

8. Changqian Yu, Changxin Gao, Jingbo Wang, Gang Yu, Chunhua Shen, Nong Sang. BiSeNet V2: Bilateral Network with Guided Aggregation for Real-time Semantic Segmentation, *International Journal of Computer Vision*, 2020, vol. 129 (11), pp. 1-18.  
DOI:10.1007/s11263-021-01515-2

9. Yingxiao Xu, Long Pan, Chun Du, Jun Li, Ning Jing, and Jiangjiang Wu. *Vision-based UAVs Aerial Image Localization: A Survey*, 2018. DOI: 10.1145/3281548.3281556

10. Efimov A.I., Novikov A.I. *Komp'yuternaya optika*, 2016, vol. 40, no. 2, pp. 258–265.

11. Ermakov P.G., Gogolev A.A. *Trudy MAI*, 2021, no. 117. URL: <https://trudymai.ru/eng/published.php?ID=156253>. DOI: 10.34759/trd-2021-117-11
12. Ivashova N.D., Mikhailin D.A., Chernyakova M.E., Shanygin S.V. *Trudy MAI*, 2019, no. 104. URL: <https://trudymai.ru/eng/published.php?ID=102223>
13. Sokolov D.Yu. *Vooruzhenie i ekonomika*, 2012, no. 1 (17), pp. 50-57.
14. Sokolov D.Yu. *Trudy MAI*, 2022, no. 123. URL: <https://trudymai.ru/eng/published.php?ID=165563>. DOI: 10.34759/trd-2022-123-17
15. Guseinov G.A., Zul'fugarly P.R., Abdurrakhmanova I.G. *Trudy MAI*, 2022, no. 126. URL: <https://trudymai.ru/eng/published.php?ID=169012>. DOI: 10.34759/trd-2022-126-26
16. Chigrinets E.G., Verchenko A.V. *Trudy MAI*, 2019, no. 104. URL: <https://trudymai.ru/eng/published.php?ID=102420>
17. Sheinikov A.A., Kovalenko A.M., San'ko A.A. *Nauchnyi vestnik MGTU GA*, 2023, vol. 26, no. 1, pp. 81-94. URL: <https://doi.org/10.26467/2079-0619-202-26-1-81-94>
18. Stepanov D.N., Tishchenko I.P. *Programmnye sistemy: teoriya i prilozheniya*, 2011, no. 4, pp. 33-43.
19. Stepanov D.N., Tishchenko I.P. *XIV natsional'naya konferentsiya po iskusstvennomu intellektu s mezhdunarodnym uchastiem, KII-2014*, Kazan', Izd-vo RITs «Shkola», 2014, vol. 3, pp. 333-341.
20. Konushin A. *Komp'yuternaya grafika i mul'timedia*, 2006, no. 4 (3). URL: <http://archive.ec/hwXHz>

Статья поступила в редакцию 03.04.2023

Одобрена после рецензирования 17.05.2023

Принята к публикации 27.06.2023

The article was submitted on 03.04.2023; approved after reviewing on 17.05.2023; accepted for publication on 27.06.2023