

Intelligent Systems and Technologies, Artificial Intelligence

Интеллектуальные системы и технологии, искусственный интеллект

Research article

DOI: <https://doi.org/10.18721/JCSTCS.19301>

UDC 004.932.72



IDENTIFICATION OF SPACECRAFT AT CLOSE RANGE USING DEPTH MAPS

*V.A. Pavlov¹ , A.K. Tsytsulin² ,
N.V. Savischenko³ , V.G. Avetisyan⁴*

¹ Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation;

² Television Research Institute, St. Petersburg, Russian Federation;

³ Military Telecommunications Academy, St. Petersburg, Russian Federation;

⁴ Russian-Armenian University (RAU), Yerevan, Armenia

✉ pavlov_va@spbstu.ru

Abstract. The article addresses the problem of automatic detection, segmentation, and classification of spacecraft using depth maps. The study was based on synthesized semi-natural data, which made it possible to simulate various orbital imaging conditions and to provide accurate ground-truth annotations for the quantitative evaluation of algorithms. Due to the instability of texture and color features under real space conditions, the geometric structure of objects is proposed as a key invariant feature. To solve the detection problem, a multi-stage algorithm has been developed, which includes adaptive histogram equalization, local threshold segmentation, morphological processing, and filtering of false contours based on geometric criteria. To classify spacecraft, deep neural networks of the ResNet family were used in two configurations: based on optical images alone and based on combined data (image + depth map). Experiments on a synthetic dataset of 19440 images showed that the proposed algorithm significantly outperforms the SAM 3 model in terms of detection accuracy (Precision 0.917 vs 0.171), ensuring a low level of false positives. The multimodal approach to classification achieved an accuracy of 99.9% (ResNet-152 model), confirming the effectiveness of using depth maps for reliable recognition of spacecraft under near-real space imaging conditions.

Keywords: spacecraft, synthetic data, object detection, computer vision, depth maps

Citation: Pavlov V.A., Tsytsulin A.K., Savischenko N.V., Avetisyan V.G. Identification of spacecraft at close range using depth maps. Computing, Telecommunications and Control, 2026, Vol. 19, No. 3, Pp. 7–17. DOI: 10.18721/JCSTCS.19301

Научная статья

DOI: <https://doi.org/10.18721/JCSTCS.19301>

УДК 004.932.72



ИДЕНТИФИКАЦИЯ КОСМИЧЕСКИХ АППАРАТОВ НА БЛИЖНЕЙ ДИСТАНЦИИ С ИСПОЛЬЗОВАНИЕМ КАРТ ГЛУБИНЫ

В.А. Павлов¹ , А.К. Цыцулин² ,
Н.В. Савищенко³ , В.Г. Аветисян⁴

¹ Санкт-Петербургский политехнический университет Петра Великого,
Санкт-Петербург, Российская Федерация;

² АО «НИИ телевидения», Санкт-Петербург, Российская Федерация;

³ Военная академия связи, Санкт-Петербург, Российская Федерация;

⁴ Российско-Армянский университет, Ереван, Армения

✉ pavlov_va@spbstu.ru

Аннотация. В статье рассматривается задача автоматического обнаружения, сегментации и классификации космических аппаратов с использованием карт глубины. Исследование проведено на основе синтезированных полунатурных данных, что позволило смоделировать различные условия орбитальной съемки и обеспечить точную эталонную разметку для количественной оценки алгоритмов. В связи с нестабильностью текстурных и цветовых признаков в реальных условиях космоса предлагается использовать геометрическую структуру объектов как ключевой инвариантный признак. Для решения задачи обнаружения разработан многоэтапный алгоритм, включающий адаптивную эквализацию гистограммы, локальную пороговую сегментацию, морфологическую обработку и фильтрацию ложных контуров по геометрическим критериям. Для классификации космических аппаратов использовались глубокие нейронные сети семейства ResNet в двух конфигурациях: на основе только оптических изображений и на основе комбинированных данных (изображение + карта глубины). Эксперименты на синтезированной выборке из 19440 изображений показали, что предложенный алгоритм значительно превосходит модель SAM 3 по точности обнаружения (Precision 0,917 против 0,171), обеспечивая низкий уровень ложных срабатываний. Мультимодальный подход к классификации позволил достичь точности 99,9% (модель ResNet-152), что подтверждает эффективность использования карт глубины для надежного распознавания космических аппаратов в условиях космической съемки, приближенной к реальной.

Ключевые слова: космические аппараты, синтетические данные, обнаружение объектов, компьютерное зрение, карты глубины

Для цитирования: Pavlov V.A., Tsytsulin A.K., Savischenko N.V., Avetisyan V.G. Identification of spacecraft at close range using depth maps // Computing, Telecommunications and Control. 2026. Т. 19, № 3. С. 7–17. DOI: 10.18721/JCSTCS.19301

Introduction

Spacecraft rendezvous control is a critical task due to the complexity of space missions. The growing number of programs for servicing stations, assembling structures in orbit, and exploring the Moon requires more precise rendezvous and docking methods. This is essential for the safety of crews in manned missions and for the success of automated missions [1].

In 1962, a method of multi-stage spacecraft rendezvous was proposed, which was a significant event in the history of space exploration. This process involved guidance based on data from Earth, semi-automatic control using radar, automatic docking at a distance of 100 to 200 meters, and final manual docking [2]. Once docking technology had been developed, television systems were used to monitor the

process. This equipment allowed the distance between the spacecraft and their relative position to be determined throughout the rendezvous phase.

In [3], an adaptive approach to spacecraft rendezvous control is proposed, which adjusts the camera's resolution and frame rate according to distance. At long distances, high resolution and low frame rate are used; at short distances, the opposite is true due to pixel grouping. This optimizes the balance between static and dynamic errors. To further optimize the system, an information function is also proposed that takes into account data loss, transmission speed, and encoder complexity. In the same work [3], a method for adapting the frame rate for monitoring spacecraft is proposed that optimizes the signal accumulation time depending on the distance to and brightness of the object. The algorithm is based on a criterion that minimizes the decision delay and the risk of error due to noise. The developed algorithm automatically controls the frame rate, which improves the efficiency of spacecraft observation and control systems.

In [4], an approach to creating a space-based video system that operates with limited *a priori* data is explored. It is based on the relationship between illumination and image dynamics, which allows the observed distance range to be expanded by adapting the frame rate.

Currently, due to the large number of spacecraft and variations in their configurations [5], an important area of research is the integration of artificial intelligence-based computer vision into onboard television systems [6].

In [7], the integration of computer vision into onboard systems for adaptive spacecraft rendezvous control is described. It is proposed to adapt the image processing and encoding parameters depending on the approach stage. This, combined with object recognition and tracking algorithms, improves control accuracy and reduces computational load.

In [8], a method is proposed for assessing the accuracy of distance and spatial orientation measurements based on the method of proportions, applicable to a single-camera matrix television system when the object type is known. When a spacecraft approaches an object under conditions of *a priori* uncertainty of the object class, the use of spatial information about the object, formed on the basis of depth maps, becomes key.

The aim of this study is to develop and evaluate a comprehensive approach for the automatic detection, segmentation, and classification of spacecraft at close range using depth maps and multimodal data to ensure reliable identification under varying orbital imaging conditions.

To achieve this aim, the following research tasks were addressed:

- 1) compiling a representative synthetic dataset of spacecraft 3D models with generated depth maps and precise ground-truth annotations;
- 2) developing a multi-stage algorithm for spacecraft detection and segmentation based on depth map analysis;
- 3) quantitatively evaluating the proposed detection algorithm against modern neural network baselines;
- 4) investigating the effectiveness of a multimodal approach to spacecraft classification by combining optical images with depth maps using ResNet architectures.

Materials and methods

In this study, 90 3D spacecraft models were selected for the experiments. The models were obtained from the “3D Resources” open database of the National Aeronautics and Space Administration (NASA)¹.

The dataset was compiled based on the principles of representativeness, technological diversity, and operational significance. The key criterion was to cover all major types of space missions. The sample includes scientific observatories for astrophysical research (such as “Hubble Space Telescope”, “Chandra X-ray Observatory”, “James Webb Space Telescope”, and others), specialized spacecraft for Earth

¹ 3D Resources – NASA Science. [online] Available at: <https://science.nasa.gov/3d-resources/> (Accessed 14.09.2026).

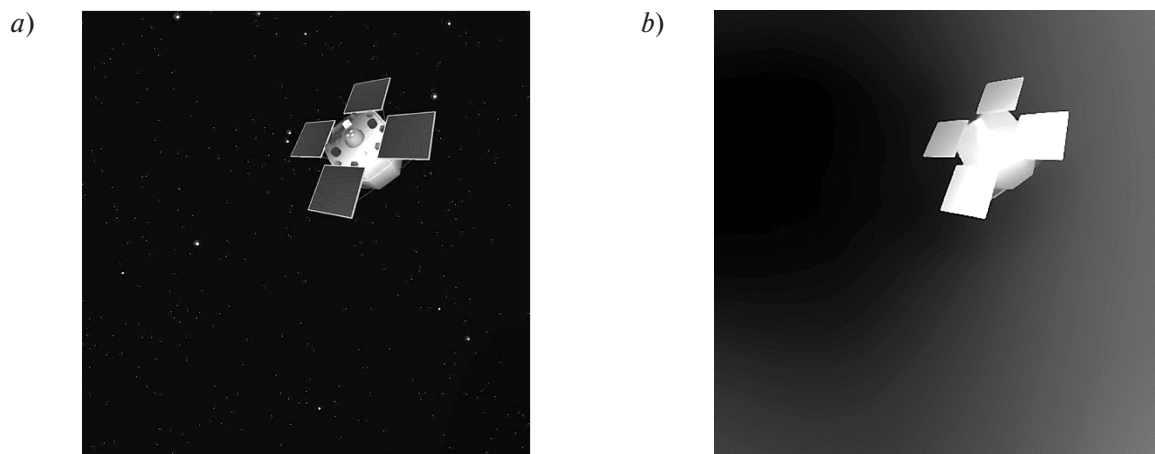


Fig. 1. Data visualization: original image of the spacecraft (a); depth map (b)

and solar observations (such as “Aqua”, “Terra”, “Solar Dynamics Observatory”, and others), and planetary and interplanetary stations (such as “Cassini-Huygens”, “Voyager”, “Dawn”, and others). To ensure the completeness of the analysis of structural forms, the set includes standardized models of small spacecraft, such as several CubeSat configurations (1RU, 2RU, ICECube etc.), which allows for comparing the unique geometric solutions of flagship missions with typical platforms. The dataset also provides historical retrospective – from early missions (“Pioneer 10”, “Skylab” etc.) to the most recent projects (“Parker Solar Probe”, “OSIRIS-Rex” etc.). In addition, it includes critical elements of space infrastructure, such as Tracking and Data Relay Satellites, which are essential for operational monitoring of near-Earth space.

Based on automatically centered 3D models, images were generated with a 60° rotation step around the Y, X, and Z axes, ensuring full coverage of all angles (216 images per class). Each image was saved on a transparent background. The objects were then placed on a realistic starry background with random scaling (from 50 to 150%) and position. The images were converted to grayscale to emphasize structural features, which is consistent with the conditions of practical space observation. For each image, a depth map was generated using Depth Anything V2 MDE model [9], which provides a relative distance to different parts of the device, preserving important 3D geometric information. These maps allow for the separation of objects from the background and the training of a classifier model based not only on visual textures but also on spatial features of the shape. Additionally, an annotation file was created for each image, containing the coordinates of the rectangle, polygon, class name, and image metadata.

Figure 1 shows a generated image of the ACRIMSAT solar observatory (Fig. 1, a) and a depth map (Fig. 1, b) obtained using Depth Anything V2 model.

Detection and segmentation of spacecraft based on depth

To solve the problem of extracting spacecraft from depth maps, an algorithm was proposed that avoids the drawbacks of neural network methods. The main criteria for development were maximum speed and autonomy: the algorithm does not require pre-training, reference databases, or the use of motion features. The key informative parameter is the depth feature, which allows for reliable segmentation of objects regardless of their dynamics. The proposed algorithm for detecting and selecting spacecraft implements a multi-stage processing of depth maps. The method is based on a step-by-step processing that includes adaptive contrast enhancement, nonlinear segmentation, morphological post-processing, and geometric filtering of false detections.

In the first stage, the depth image is pre-processed using Contrast Limited Adaptive Histogram Equalization (CLAHE) [10]. The CLAHE algorithm splits the image into a regular grid of 8×8 pixel

blocks and performs histogram equalization in each block independently, limiting the contrast gain factor to C_{Limit} . Mathematically, the operation is described as:

$$I_e(x, y) = CLAHE(I_{depth}(x, y), S_{grid}, C_{Limit}),$$

where I_{depth} is the original depth map.

Adaptive thresholding binarization [11] is used for object segmentation, which calculates a local threshold $T(x, y)$ for each pixel based on the average intensity value in a 11×11 pixel neighborhood:

$$B(x, y) = \begin{cases} 255, & I_e(x, y) > \mu_{local}(x, y) - K \\ 0, & \text{else} \end{cases},$$

where μ_{local} is the local mean in the Gaussian window, and K is the offset that provides sensitivity to low-contrast depth gradients.

The binary mask often contains noise and discontinuities. To eliminate them, the morphological closing operation is used:

$$M_{closed} = (B \oplus S) \ominus S,$$

where $\oplus$ and $\ominus$ are dilation and erosion operations, respectively, and S is a 5×5 structural element (kernel). This procedure eliminates small holes inside objects and connects closely spaced fragments. A subsequent dilation step compensates for gaps in the contours of elongated structures, such as solar panels and satellite antennas.

Contour detection is implemented using the algorithm described in [12], which builds a hierarchy of contours using boundary topology analysis. To eliminate object fragmentation, a heuristic algorithm is used to merge nearby contours based on the distance between their centers of mass. The center of mass of a contour C is calculated using the moments:

$$c_x = \frac{M_{10}}{M_{00}}, \quad c_y = \frac{M_{01}}{M_{00}},$$

where M_{00} is the zero moment (area), and M_{10} and M_{01} are the first moments.

False detection filtering is based on geometric shape invariants. Each contour is estimated by its area A , perimeter P , and circularity coefficient:

$$C = \frac{4\pi A}{P^2}.$$

The remaining contours are sorted by area in descending order, and no more than two largest objects are retained. For each object, a bounding rectangle R is constructed and projected onto the original image for visualization.

Standard metrics based on the intersection of predicted and true bounding boxes were used to quantify the effectiveness of the detection algorithm [13]. The basic formulas for calculating key indicators are presented below.

Intersection over Union (IoU) is a measure that determines the degree of overlap between the predicted B_p and the true B_{gt} rectangle:

$$\text{IoU}(B_p, B_{gt}) = \frac{\text{Area}(B_p \cap B_{gt})}{\text{Area}(B_p \cup B_{gt})}.$$

The IoU threshold (usually denoted as θ) is used to classify a prediction as either a true positive (TP) or a false positive (FP).

Precision is the proportion of correct detections among all predicted detections:

$$\text{Precision} = \frac{TP}{TP + FP}.$$

Recall is the proportion of true objects detected among all objects in the sample:

$$\text{Recall} = \frac{TP}{TP + FN},$$

where FN (False Negative) is the number of undetected true objects.

F1-score is the harmonic mean between Precision and Recall:

$$\text{F1-score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}.$$

To evaluate the quality of semantic image segmentation, the Dice coefficient is used, which measures the degree of intersection between the predicted M_p mask and the M_{gt} reference:

$$\text{Dice}(M_p, M_{gt}) = \frac{2 \times |M_p \cap M_{gt}|}{|M_p| + |M_{gt}|},$$

where $|M_p \cap M_{gt}|$ is the number of common pixels in both masks (intersection), $|M_p| + |M_{gt}|$ is the total number of pixels in the predicted and reference masks, respectively.

The Segment Anything Model 3 (SAM 3) [14] was chosen as the modern baseline approach for comparison, due to its ability to operate in “zero-shot” mode without requiring pre-training or fine-tuning on target data. Table 1 presents the test results on a sample of 19440 images (90 classes $\times$ 216 images per class).

Table 1

Testing results of the proposed algorithm and SAM 3 on a test sample

Method	Precision	Recall	F1-score	IoU	Dice
Proposed	0.917	0.968	0.942	0.877	0.853
SAM 3	0.171	0.342	0.228	0.841	0.909

Figure 2 shows an example of detecting (a) and segmenting (b) the ACRIMSAT spacecraft using depth maps.

As can be seen from the Table 1, the proposed algorithm significantly outperforms the baseline SAM 3 in terms of the main detection metrics. The Precision of the proposed method is 0.917, which is more than five times higher than the corresponding value of SAM 3 (0.171). This indicates an extremely low

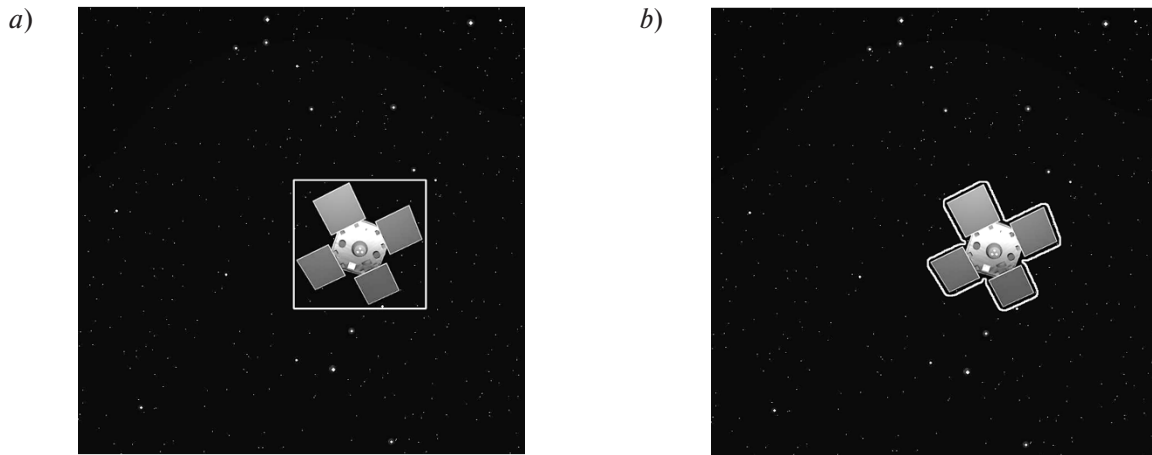


Fig. 2. Example of detection (a) and segmentation (b) of the spacecraft

level of false positives: there were only 1705 false positives compared to 32177 for SAM 3. The high Recall value (0.968 vs 0.342) confirms the reliability of object detection: the algorithm successfully identified 18824 out of 19440 objects (96.83%), missing only 616 (false negative).

Despite the fact that the SAM 3 demonstrates segmentation quality indicators comparable to those of a specialized algorithm (Dice 0.909 and Mask IoU 0.841), its practical application for the task of spacecraft detection is ineffective due to the extremely high level of false positives. The low Precision value of SAM 3 is due to its architecture, which is designed for the segmentation of arbitrary objects, which in the context of specific data leads to numerous false detections against the background of stars, noise, or the spacecraft's own structural elements. The proposed algorithm, due to the use of depth features and multi-stage geometric filtering, provides high detection accuracy.

A quantitative assessment of the segmentation quality of the proposed algorithm (average Dice = 0.853, average IoU = 0.877) confirms that the generated masks are well aligned with the true object boundaries. The slightly lower Dice value compared to SAM 3 is compensated by higher localization accuracy and the absence of training requirements. Therefore, the developed method provides an optimal balance between the completeness of detection and the accuracy of contour extraction, making it more suitable for autonomous depth image processing systems.

Classification of spacecraft

In the conditions of orbital imaging, textural and color features are often unstable, while the geometric structure of the object remains invariant to the conditions of external illumination. In this regard, the use of multimodal approaches that combine semantic information from the image with information about the spatial structure of the object becomes particularly relevant.

ResNet architectures [15] were chosen for the classification of spacecraft, as their structural features allow for effective solution of this computer vision tasks. The key factor in the choice was the concept of residual learning, implemented through blocks with through-the-stack connections. This architectural solution avoids the problem of gradient decay, which is critical when training deep neural networks. Images of spacecraft are characterized by a high variety of angles, specific illumination conditions in outer space, and the presence of many small structural elements. The ResNet architecture provides deep feature differentiation: the initial layers of the network successfully identify the boundaries and textures of surfaces, while the deeper layers form a high-level representation of the topology of specific types of spacecrafts.

The experiment was conducted using ResNet architectures (18, 34, 50, 101, 152) pre-trained on the ImageNet dataset. The output fully connected layer of each model was replaced with a new linear

Table 2

Training parameters for classifiers

Parameter	Value
Batch size	64
Optimization method	Adam
Initial learning rate	10^{-3}
Loss function	Cross-entropy
Number of epochs	30

layer with the number of neurons equal to the number of spacecraft classes (90 classes). The training parameters are presented in Table 2.

All input images were scaled to a size of 128×128 pixels. Each model was trained on an Nvidia RTX 2080 TI graphics accelerator with 11 GB of video memory.

Based on the generated scenes with spacecraft, a dataset was prepared for solving the classification task. From each pair of images and depth maps, the area containing the spacecraft was extracted using annotations. The dataset was randomly divided into three non-overlapping sets: training set (70% of the dataset, 13680 images), validation set (15% of the dataset, 2880 images), and test set (15% of the dataset, 2880 images).

Two series of experiments were conducted to evaluate the effectiveness of ResNet. In the first stage, the models were trained exclusively on the original spacecraft images. A comparative analysis of the quality metrics of the ResNet models on spacecraft images is presented in Table 3.

Table 3

Comparative analysis of ResNet model quality metrics on spacecraft images

Model name	Layers number	Accuracy	Precision	Recall	F1-score
ResNet18	18	0.964	0.964	0.964	0.964
ResNet34	34	0.967	0.965	0.965	0.965
ResNet50	50	0.970	0.971	0.971	0.971
ResNet101	101	0.976	0.978	0.977	0.977
ResNet152	152	0.979	0.980	0.979	0.979

The results presented in Table 2 demonstrate an increase in classification quality with an increase in the depth of the architecture. The ResNet18 model shows high accuracy (0.964), which confirms the effectiveness of residual connections even for relatively shallow networks. The transition to ResNet34 and ResNet50 provides an Accuracy increase of 0.3% and 0.6%, respectively, due to a richer hierarchy of extracted features. The best result among models trained on visual images only is demonstrated by ResNet152 (Accuracy 0.979). This shows that deep architecture is capable of capturing subtle differences in textures and shapes of 90 classes of spacecraft. All metrics (Precision, Recall, F1-score) are balanced and almost identical to Accuracy, which indicates that there is no class imbalance in the test set and that the predictions are stable.

The key result of the study was obtained in the second series of experiments, where the classification of combined data (an image supplemented with a depth map) was performed. A comparative analysis of the quality metrics of the ResNet models on combined images is presented in Table 4.

Table 4

Comparative analysis of ResNet model quality metrics on combined images

Model name	Layers number	Accuracy	Precision	Recall	F1-score
ResNet18	18	0.988	0.988	0.987	0.988
ResNet34	34	0.992	0.993	0.992	0.993
ResNet50	50	0.996	0.996	0.996	0.996
ResNet101	101	0.998	0.999	0.999	0.999
ResNet152	152	0.999	0.999	0.999	0.999

Adding depth maps to the original images improves all quality metrics for each of the architectures considered. The baseline ResNet18 model on the combined data (Accuracy 0.988) outperforms the deepest single-modal ResNet152 model (Accuracy 0.979). This strongly suggests that geometric information (depth) is a more stable and informative feature in space imaging conditions than texture or color.

The ResNet34 architecture has the highest relative increase (Accuracy increased from 0.967 to 0.992, an increase of +2.5%). The deepest models – ResNet101 and ResNet152 – achieve almost perfect results (Accuracy 0.998 and 0.999, respectively). The classification error of 0.1% for ResNet152 can be explained by the presence of heavily noisy or non-standard viewing angles, where even geometric information is interpreted ambiguously. Nevertheless, the 99.9% result confirms that the “image + depth map” combination allows ResNet-based networks to solve the task of spacecraft classification with high reliability, minimizing the impact of external conditions and visual interference.

Conclusion

As part of this work, based on synthesized semi-natural data, a complete data processing cycle for automatic spacecraft recognition was developed and tested, including the stages of detection, segmentation and classification. The use of synthesized scenes made it possible to provide accurate reference markings (ground truth) and controlled modeling of various orbital survey conditions, which is important for quantifying the proposed algorithms.

To solve the detection problem, an algorithm based on the analysis of depth maps is proposed. Unlike neural network approaches, it does not require pre-training and demonstrates high efficiency. An experimental comparison with the modern neural network model SAM 3 on a synthetic dataset showed that the proposed algorithm provides superior detection accuracy (Precision 0.917 vs 0.171) while maintaining high completeness (Recall 0.968).

In the classification task, the potential of a multimodal approach that combines standard images and depth maps was explored. A series of experiments with ResNet-based architectures confirmed that adding depth information improves recognition accuracy. Combining images and depth maps allowed the ResNet152 model to achieve high classification rates (Accuracy of 99.9%), while classification based solely on optical images achieved around 97.9% accuracy.

Consequently, the research tasks formulated in the introduction have been fully accomplished. The developed comprehensive approach, which combines fast depth-based detection and accurate multimodal classification, is proven to be an effective and reliable solution for autonomous navigation, inspection, and tracking of space objects at close range, adapting to varying orbital conditions, fully achieving the stated aim of the study.

REFERENCES

1. **Sagdullaev Iu.S., Smirnov A.I., Kovin S.D.** *Osnovy televizionnykh izmerenii parametrov sbliizheniia kosmicheskikh apparatov* [Fundamentals of Television-Based Measurement of Spacecraft Approach Parameters], monograph, Moscow: Sputnik+, 2014
2. **Keldysh M.V.** *Tvorcheskoe nasledie akademika Sergeia Pavlovicha Koroleva* [Creative Legacy of Academician Sergei Pavlovich Korolev], Moscow: Izdatel'stvo "Nauka", 1980.
3. **Devyatkin A.V., Gorshanov D.L., Tsytulin A.K., Morozov A.V., Bobrovsky A.I., Pavlov V.A.** Adaptation of frame frequency to observation stages at control of spacecraft convergence. *Journal of Physics: Conference Series*, 2019, Vol. 1236, Art. No. 012069. DOI: 10.1088/1742-6596/1236/1/012069
4. **Tsytulin A.K., Bobrovskiy A.I., Morozov A.V.** Synthesis of a space videoinformation system with a wide range of changes in the distance to the object. *Voprosy radioelektroniki. Seriya: tekhnika televideniia* [Issues of Radio Electronics. Series: Television Technology], 2020, Vol. 2, Pp. 19–37.
5. **Klyushnikov V.Yu.** Construction of small spacecraft clusters. *Izvestia vuzov. Priborostroenie*, 2016, Vol. 59, No. 6, Pp. 423–428. DOI: 10.17586/0021-3454-2016-59-6-423-428
6. **Balukhto A.N., Romanov A.A.** Artificial Intelligence in Space Technology: State, Development Prospects. *Rocket-Space Device Engineering and Information Systems*, 2019, Vol. 6, No. 1, Pp. 65–75. DOI: 10.30894/issn2409-0239.2019.6.1.65.75
7. **Bobrovskiy A.** Video processing in adaptive control system of spacecraft rendezvous. *Information and Space*, 2018, Vol. 4, Pp. 149–153.
8. **Bachevskiy S.V.** Tochnost' opredeleniia dal'nosti i orientatsii ob'ekta metodom proporsii v matrichnykh televizionnykh sistemakh [Accuracy of definition of range and orientation of object by the method of proportions in matrix television systems]. *Voprosy radioelektroniki. Seriya: tekhnika televideniia* [Issues of Radio Electronics. Series: Television Technology], 2010, No. 1, Pp. 57–66.
9. **Lihe Yang, Bingyi Kang, Zilong Huang, Zhen Zhao, Xiaogang Xu, Jiashi Feng, Hengshuang Zhao.** Depth Anything V2. *arXiv:2406.09414*, 2024.
10. Zuiderveld K. Contrast Limited Adaptive Histogram Equalization. In: *Graphics Gems IV* (ed. P.S. Heckbert), 1994, Pp. 474–485.
11. **Bradley D., Roth G.** Adaptive thresholding using the integral image. *Journal of Graphics Tools*, 2007, Vol. 12, No. 2, Pp. 13–21. DOI: 10.1080/2151237X.2007.10129236
12. **Suzuki S., Abe K.** Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 1985, Vol. 30, No. 1, Pp. 32–46. DOI: 10.1016/0734-189X(85)90016-7
13. **Everingham M., Van Gool L., Williams C.K.I., Winn J., Zisserman A.** The PASCAL Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 2010, Vol. 88, Pp. 303–338. DOI: 10.1007/s11263-009-0275-4
14. **Carion N., Gustafson L., Hu Y.-T. et al.** SAM 3: Segment Anything with Concepts. *arXiv:2511.16719*, 2025. DOI: 10.48550/arXiv.2511.16719
15. **Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun.** Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, Pp. 770–778. DOI: 10.1109/CVPR.2016.90

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ**Vitalii A. Pavlov****Павлов Виталий Александрович**

E-mail: pavlov_va@spbstu.ru

ORCID: <https://orcid.org/0000-0003-0726-6613>**Aleksandr K. Tsytulin****Цыцулин Александр Константинович**

E-mail: tsytulin@niitv.ru

ORCID: <https://orcid.org/0009-0008-1065-6798>**Nikolay V. Savischenko****Савищенко Николай Васильевич**

E-mail: snikaspb@mail.ru

ORCID: <https://orcid.org/0000-0001-5282-2506>**Vahan G. Avetisyan****Аветисян Ваан Генрихович**

E-mail: avahan@mail.ru

*Submitted: 20.03.2026; Approved: 20.07.2026; Accepted: 28.07.2026.**Поступила: 20.03.2026; Одобрена: 20.07.2026; Принята: 28.07.2026.*