Autonomous systems that belong to the field of artificial intelligence applications will become increasingly important for technical system developers in the future. The tasks of interaction on heterogeneous platforms of autonomous technical systems are among the priority areas of investment in many countries. Such technical complexes find various applications, for example, for filming various areas, inspection of bases, and dungeons, reconnaissance with semi-autonomous navigation. Having the capability to autonomously plan, reconfigure, and perform tasks for an unmanned aerial vehicle (UAV) is a crucial component of UAVs group, enabling the exploration of large, complex, and unknown environments, like in these competitions [1]. Also, it is important to develop innovation effectively and to provide solutions to the problems that matter to roboticists and society. It becomes necessary to have development platforms that facilitate the acquisition of basic competences and skills for programming video processing methods using low-cost small UAV models.

The research is targeted at the choosing a way of structuring visual information for the autonomous flight of a low-cost small UAV; then simulate and perform a test flight over a given area indoors with the construction of a map of the detected obstacles. The object of the study is the monitoring of indoor processes using video information obtained by an autonomous UAV. The subject of the research is the methods and means of processing video information received from a UAV video camera, using semantic data segmentation technologies.

Keywords: unmanned aerial vehicles; convolutional neural network; semantic segmentation; flight control system; occupancy grid.
Structurally, the information is presented as follows: in the first section, the existing hardware platforms for autonomous flight were analysed. In the second section, the methods of structuring visual information were analysed, and in the third section, the strategies and choice of a flight research scenario in unknown spaces was analysed. In the fourth section, the developed method and results in the form of constructed barrier maps was presented. To develop a method of autonomous UAV flight based on video camera data, the semantic segmentation network learning method, computer algorithms in probabilistic robotics for mobile robots were used.

1. Analysis of existing hardware platforms and scenarios for autonomous indoor flight

UAV autonomous operation modelling is practiced on micromodels, among which the most popular are: PARROT Mambo model, DJI Tello EDU, DJI ROBOMASTER TT, and Bitcraze Crazyflie. The hardware platform of such UAVs allows for autonomous flight indoors. To work with these UAV models are provided: a software development kit for Parrot Copters [2]; Tello SDK 2.0 [3]; ROBOMASTER TT SDK 3.0 [4]; open libraries for copter Crazyflie [5]. Well-known programming environment MATLAB contains packages to simulate these UAVs [6, 7]. DJI UAVs models have more limitations when planning an autonomous flight since the control systems are closed for modifications, and very popular tools such as the Bitcraze Crazyflie UAV model cannot be reprogrammed with high-level languages. Therefore, the PARROT model was chosen for further consideration. MATLAB platform and built-in UAV PARROT tool has a high degree of connectivity [8], and integrated simulation based on hardware allows you to perform a quick and realistic test of both aircraft and flight scheduling algorithms.

When performing tasks with autonomous UAVs, it is important to determine the purpose of the flight to set algorithms for the behavior of the UAV in real time, considering the characteristics of the environment and of the tasks to be executed. It is also necessary to implement some form of accelerated task planning, in case of perturbations, or changes in environmental conditions. Three types of tool kits have been applied in such scenarios: localization, mapping, and planning. The localization kit uses onboard sensor information, such as that provided by a stereo camera [8, 9]. The used architecture gives an accurate localization result (error < 0.3 m). It allows you to obtain a cloud of points, which can later be used in the system for fly planning. But this subsystem is not available for low-cost small UAV [10]. To solve such problems, the analysis of visual information coming from cameras installed on board UAVs is often used [11]. In studies [9, 11], localization is carried out in conjunction with flight planning. Here, the flight must occur in a prepared room where the specified markers are located, usually of a contrasting colour. In work [9], a visual toolkit was proposed for flight planning. This is important for teaching visual navigation techniques on inexpensive UAV models. In [10] are proposed collision-free path planner based on the rapidly exploring random trees variant, for safe and optimal navigation of robots in 3D spaces. When developing an alternative that works only on camera video data, it is necessary to compare the results with such models.

Theoretical research on this problem is aimed at developing methods for structuring visual information [11, 12] and its transformation into data and commands to change the behavior of UAVs. When implementing these methods in the practice of UAV flights, it is important to process the input data: eliminate noise in the obtained images, consider external flight factors (wind, lighting), and consider hardware errors.

2. Methods for structuring visual information of UAV PARROT

Smart UAV technologies include two main areas of research. The first area includes the development of intelligent flight controllers and path planning methods. The second area focuses on applying deep learning techniques to extract useful information from sensory data collected by the drone [13]. For autonomous UAVs, these areas of study are often considered together because after extracting useful information (for example, from a video), it is necessary to make real-time decisions (e.g., path or task planning) based on such data.

One very straightforward option is to use computer vision has the main or unique source of information, and perform autonomous navigation accordingly [14, 15]. This is also the scenario investigated in this paper. For example, consider the movement of a UAV over a curved line drawn on the floor. The proposed model consists of two parts (fig.1): analysis of video stream image data and operation of the control system. In this task, it is important to recognize a line that is in contrast with other flooring elements [16]. In case of corners appearing, it is necessary to change the yaw angle, and in case there is no line on the video stream image, the UAV must find the landing marker (circle) and land.

Reference [17] suggests deep learning as a tool to develop a vision-based UAVs Pursuit-Evasion. A deep convolutional neural network (CNN) is used to detect objects of interest (UAV) and estimate the necessary controls for the follower UAV to keep the target UAV within its field of view and the closest possible to the centre of the image frame.
YOLO v2 was used as the UAV detector since it was the best performing in complex outdoor conditions and faster enough to enable the processing at a rate of 30fps for a real-time tracking of the UAV. Deep learning and CNN are also used to train agents that control mini UAVs based on hand gestures in [18].

Following the example of implementing semantic segmentation for terrestrial autonomous vehicles [19], this method of extracting data from video is used to control the movement of UAVs [20, 21]. The work [22] provides a link for applying a trained semantic network for UAV, which is used outside. The main semantic segmentation algorithms which are used for UAV video data are: grayscale image processing, conditional random field, and deep learning. The datasets for image segmentation are used as input data for setting up the semantic segmentation system and subsequent decision making about the flight plan based on video data. After training the system, it is necessary to check the correctness of its operation. The most popular metrics are Pixel Accuracy, Mean Pixel Accuracy (mPA), Intersection over Union (IoU), Jaccard index, Dice index, and F1-score [23].

The practical result of this work is the creation of an obstacle avoidance system for UAVs using only a monocular camera (available in low-cost small UAVs). In [22] used the feature point detector Speeded Up Robust Features for fast processing of obstacles, on unknown positions. Extended StixelWorld [20] used colour information to learn the model’s obstacles. Deep neural network models have recently demonstrated remarkable performance improvements shown to outperform most traditional methods. Also, vision-based methods have poor performance under extreme illumination conditions such as shadows and direct sunlight.

The choice of the UAV visual information structuring method significantly affects the amount of data received, and therefore the speed of obtaining it. Accordingly, a specialized UAV control system will be algorithmically connected to the visual information processing subsystem.

3. Scenarios of autonomous flight of UAV PARROT using visual data

Consider the scenarios that are used by UAVs for autonomous flight, considering visual information indoors:

- patrolling the environment. In this scenario, the dimensions of the environment are known in advance. Based on the data on the size of the premises and the technical capabilities of the UAV, the route (snake, chaotic, etc.), landing conditions are selected. Visual information can be transmitted during the flight or analysed after landing. In the Parrot Mambo model, it is possible to monitor from the lower camera, originally built into the UAV body, but it is intended for navigation. For live broadcasting, an additional camera equipped with a transmitter is used. Reviews indicated that in the case of Mambo PARROTS, the transmitted image lags behind the actual camera view [2];

- recognition and observation of the object (fig. 2). Here, an object must be known in advance, which can be stationary (for example, a line drawn on the floor) or moving (for example, people or another UAV). Preliminary training or adjustment of the video information processing subsystem for this object is required. During the flight, the UAV analyses the features of a given object in the focus of the camera, after recognition, as a rule, its location is estimated and algorithms for changing the coordinates, speeds, or angles of the UAV flight are performed;

- exploration of an initially unknown area indoors (fig. 3). In such a case, it is necessary to have semantic information from a deep learning model. Based on these data, a virtual map is created during the flight. Since we assume that the PARROT UAV flies at a certain altitude, and due to the limited computing capabilities of the video information processing subsystems and the control system, we assume that the mapping will be restricted to 2D maps.
Fig. 2. Recognition and observation of the object

![Image](image1.png)

*a) start fly  b) fly around the territory*

**Fig. 3. Virtual 2D map**

The map consists of minimal square blocks (block side length l). The choice of block size is related to the UAV movement algorithm. Fig. 3 shows an example of how the system works. The map contains free Vfree, occupied Voc, and unknown Vun areas for flights.

4. The method of exploration of an initially unknown area indoors

The PARROT UAVs are equipped with a FPV camera that provides images measuring 640x360 pixels. The image data are used to develop vision-based algorithms. Therefore, the image data obtained from the FPV camera is a 360-by-640-by-3 matrix of type uint8, in RGB format.

For simulation purposes, it is important to define how to represent the camera «Field of View» (FOV). The camera will be attached at the centre of UAV recording toward its X-Axis body frame. To be able to represent the camera FOV, we will need the following camera specifications:

- depth of view D, which represents the maximum distance the camera can clearly record D=2 m;
- the angle in which the camera lens can record (θ = 110°).

These parameters give an approximate coverage of the considered area of 1 m × 2 m. To prepare the data and the autonomous flight algorithm, the UAV made a test with video recording. The resulting video was processed in the subsystem pixelLabelTrainingData (c). As a result, data were obtained (“Image datastore” and “Pixel label datastore”) for training the semantic segmentation network (the function trainNetwork was used).

The resulting semantic segmentation network (SSN) will be used to automate the flight with the following conditions:

- the flight will be made above the floor,
- when recognizing furniture objects, it is necessary to estimate the distance to them, for subsequent placement on a virtual map.

The next important task is to determine the distance to the detected furniture objects. To calculate the distance, you should know the internal and external parameters of the camera. The internal parameters of the camera can be found using the calibration procedure in (Single Camera Calibrator App) and the external parameters:

- the height of its placement above the floor (UAV flight height – h, fig. 4),
- tilt angle (depends on the pitch angle).

To find the distance to the object on the segmented frame, the coordinates of the rectangular area (bbox) describing the UAV are found, and the coordinates of the point t1 are calculated, and then, using the coordinate transformation, the coordinates of the point t1 are calculated in the “top view” coordinate system [24]. In the new system, the x coordinate is the distance to the opposite object (L).
The resulting trained SSN model and determining the distance to obstacles method are used in the UAV flight algorithm.

The method consists of the following steps:

1. The virtual space matrix Occupancy grid is initialized. The UAV is taking off, and the cell in which the UAV is taking off is considered the Vfree area, and the rest are Vun.

2. The UAV makes a 360° turn, after which 4 image data corresponding to 0°, 90°, 180°, 270° are processed. In each image, the number of pixels corresponding to the "floor" category is calculated. The results were ranked and recorded in the priority direction matrix.

3. If UAV movement direction, with the highest priority belongs to the Vfree category:

   3.1. The UAV turns to an appropriate angle.
   3.2. The distance L to the obstacle is estimated.
   3.3. Cells that are at a distance L are fixed as Vocc.
   3.4. The UAV flies the distance L/2 and stops, each flight cell is fixed as Vfree in the matrix Occupancy grid, then step 2 is performed (fig. 5).

   Else: if the directions of movement are not completed: the next priority is chosen, else the completion of the UAV flight.

The occupancy grid is a grid of values, each of which indicates an obstacle in a specified area. Values can be binary (0 for empty cells, 1 for cells occupied by an obstacle) or take values in a given range, indicating the possibility of the specified area.

On fig. 5, the height of the obstacles is set randomly. Thus, the map considers a room 4×5 m with 5 obstacles, which are pieces of furniture.

The quality of the proposed method depends on the algorithm of semantic segmentation, which is used at the second stage. The standard measure used to evaluate the performance of semantic segmentation’s algorithms is the IoU. Given an image, the IoU measure gives the similarity between the predicted region and the ground-truth region for an object present in the image, and is defined as the size of the intersection divided by the union of the two regions [23]:

$$\text{IoU}_c = \frac{TP_c}{TP_c + FP_c + FN_c},$$

where TP_c, FP_c, FN_c denote the number of true-positive, false-positive, and false-negative pixels, respectively for class c. After evaluating the recognition results of individual classes, we found the average IoU (Tabl. 1).

<table>
<thead>
<tr>
<th>IoU measure</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>floor</td>
<td>0.93</td>
</tr>
<tr>
<td>furniture objects</td>
<td>0.47</td>
</tr>
<tr>
<td>walls</td>
<td>0.88</td>
</tr>
<tr>
<td>windows</td>
<td>0.24</td>
</tr>
<tr>
<td>Averaged IoU</td>
<td>0.63</td>
</tr>
</tbody>
</table>

The video stream data received after the test flight was segmented. The segmentation results were compared with manual segmentation. A result IoU>0.5 is considered acceptable.
5. Discussion

Clearly, some limitations have to be considered, which include the battery autonomy (about 10 minutes) and the computational power onboard of the UAV. These limitations can be considered, however, also a means to challenge the creativity of the users. With regard to this aspect, this paper only presents some of the preliminary ideas and their margins for improvement. We have shown that the hardware is suitable for performing a phased flight of an unknown area based on the segmentation of the environment. Using it, as well as the method of obtaining the distance to objects-obstacles of the MATLAB library, we have successfully developed an algorithm for the autonomous flight of a small UAV. In the future, it is necessary to conduct a detailed campaign of experiments to assess the benefits and limitations of the semantic segmentation and develop more robust (less sensitive to parameter calibration) and adaptive autonomous planning and navigation algorithms.

The specific drawbacks that should be taken into account are mainly related to the sensing devices used, that is, the monocular camera has the drawback of the high sensitivity to lighting conditions; such as direct sun light may led to a lack of information.

Conclusions

The autonomous flight method of UAV Parot Mambo using semantic segmentation data for objects indoors is developed in the article.

The scientific novelty of the study lies in a was improved the method of autonomous flight of small UAVs by using the semantic network model and determining the purpose of flight only at a given altitude to minimize the computational costs of limited autopilot capabilities for low-cost small UAV models.

The study's practical significance lies in the fact that we design a model-based algorithm on Simulink and Matlab through simulation, and study how to test it by deploying on the Hardware of a Parrot Mambo Fly minidrone through the interfaces provided with its Hardware support package in Simulink. We used the generated source code in real-time applications such as rapid prototyping, simulation, and hardware-in-the-loop tests with Simulink Coder.

During the test flight, the results of the semantic segmentation were displayed on the screen to assess the quality of the resulting neural network. In parallel, statistics were collected to calculate the IoU. Because of several flights, IoU=0.63 was obtained, which is an acceptable result for further application in UAV automatic flight algorithms.

Further research will be aimed at solving the problem of occupancy map construction accuracy and increasing UAV flight time. We are planning changes in the UAV hardware architecture because the coordinates of obstacles will also be obtained from additional devices: lidar, laser pointer. It will make it possible to specify the coordinates of both the UAV position and the position of surrounding objects. The software architecture will be changed too. For the final calculation of the coordinates provided by various sensors, the capabilities of the open packages ROS 2 and the NAV2 tool will be used. The planned decision is to maintain the capability to hardware-in-loop testing of the actual flight control system. This will significantly speed up the developing of a flight control system.

Contributions of authors: formulation of the purpose and tasks of research, formulation of conclusions – David Nasso; analysis of methods for structuring visual information – Rossella Bartolo; analysis of scenarios of autonomous flight – Sergiy Yashin; development of methods and analysis of research results – Olha Pohudina; evaluation of the quality of the semantic segmentation algorithm – Andrii Pohudin.

All authors have read and agreed to the published version of the manuscript.

Acknowledgement. The work was supported by the Ministry of Education and Science of Ukraine, and the state registration numbers of the projects are 0121U109605 and 0122U000860.

References

5. Giernacki, W., Skwierczyński, M. et al. Crazyflie 2.0 quadrotor as a platform for research and education in robotics and control engineering. 22nd International Conference on Methods and Models in Automation and
МЕТОД ЗАБЕЗПЕЧЕННЯ АВТОНОМОГО ПОЛЬЗОВУ БЕЗПІЛОТОНОГО ЛІТАЛЬНОГО АПАРАТУ PARROT МАМБО З ВИКОРИСТАННЯМ ДАНИХ СЕМАНТИЧНОЇ СЕГМЕНТАЦІЇ

Давід Насо, Ольга Погудіна, Сергій Яшин,
Андрій Погудін, Россела Бартоло

За останнє десятиліття автономна навігація безпілотних літальних апаратів (БПЛА) є надзвичайно прийнятною темою, у тому числі через підвищення доступності обладнання та програмного забезпечення з відкритим вихідним кодом для керування та обробки даних. Попит на дані технології викликає великий інтерес...
до розробки доступних експериментальних платформ для навчання студентів інженерних спеціальностей в швидко розвиваючій галузі автономної навігації. У цій статті ми описуємо платформу, що базується на недорогому обладнанні, яка використовує переваги недорогих моделей маленьких БПЛА за рахунок використання семантичної сегментації. Результатом нашого дослідження є метод, який підходить як для забезпечення автономного польоту БПЛА із фронтальною відеокамерою. Наукова новизна одержаних результатів полягає в наступному: удосконалено метод автономного польоту малих БПЛА за рахунок використання семантичної мережевої моделі та визначення мети польоту лише на заданій висоті для мінімізації обчислювальних витрат обмежених можливостей автопілоту. Результати нашого дослідження можуть бути розширени шляхом проведення серії експериментів для різних умов.

**Ключові слова:** Безпілотні літальні апарати; змістовно-обчислювальна среда; декілька обмежень визначення мети польоту для автономного польоту БПЛА через несивідомі умови, формалізування процедур інформаційної втрати;

Давід Насо – PhD, проф. зав. лаб. робототехніки кафедри електротехніки та інформаційної техніки, Політехнічний університет Барі, Італія.

Ольга Погудіна – канд. техн. наук, доц., доц. каф. інформаційних технологій проектування, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, Україна.

Сергій Яшин – стар. наук. співроб. каф. інформаційних технологій проектування, Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут».

Андрій Погудін – PhD, доц. каф. космічної техніки та нетрадиційних джерел енергії Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут».

Росселя Бартоло – PhD, наук. співроб. каф. механіки, математики, менеджменту, Політехнічний університет Барі, Італія.

David Naso – PhD, Full Professor, Head of the Robotics Lab at the Department of Electrical and Information Engineering, Politecnico di Bari, Italy, e-mail: david.naso@poliba.it, ORCID: 0000-0002-7948-8226, Scopus Author ID: 6602866170.

Olha Pohudina – Candidate of Technical Sciences, Assistant Professor of the Department of Information Technology Design, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine, e-mail: ok.pogudina@gmail.com, ORCID: 0000-0001-5689-2552, Scopus Author ID: 57204907264.

Andrii Pohudin – PhD, Assistant Professor of the Department of Space Technology and Non-Traditional Energy Sources, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine, e-mail: andrey.pohudin@gmail.com, ORCID: 0000-0002-8673-2119, Scopus Author ID: 57214223045.

Sergey Yashin – senior researcher of the Department of Information Technology Design, National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, Ukraine, e-mail: sergey.yashin.xai@gmail.com, ORCID: 0000-0002-1680-4156, Scopus Author ID: 57219053677.

Rossella Bartolo – Ph. D. cum laude in Mathematics (Geometry and Topology Department), Researcher of the Department of Mechanics, Mathematics, Management, Politecnico di Bari, Italy, e-mail: rossella.bartolo@poliba.it, ORCID: 0000-0002-7230-9651, Scopus Author ID: 7004507482