http://doi.org/10.35784/iapgos.7090 *received: 15.01.2025 | revised: 06.04.2025 | accepted: 21.04.2025 | available online: 27.06.2025*

OBJECT DETECTION ALGORITHM IN A NAVIGATION SYSTEM FOR A RESCUE DRONE

Nataliia Stelmakh[1], Yurii Yukhymenko[1], Ilya Rudkovskiy[1], Anton Lavrinenkov[2]
[1]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Department of Computer-Integrated Technologies of Device Production, Faculty of Instrumentation Engineering, Kyiv, Ukraine, [2]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Department of Aircraft Manufacturing Engineering of the Education Research Institute of Mechanical Engineering, Kyiv, Ukraine

Abstract. *This article focuses on improving object recognition algorithms for rescue drones, in particular, enhancing the methodology for classifying human poses by expanding the set of key body points and using more effective mathematical models. A methodology is proposed that analyzes 11 key body points, enabling the classification of human positions (standing, lying down, sitting, kneeling, bent) with greater accuracy. Additionally, a gesture recognition algorithm is proposed, detecting gestures such as arm-waving as a signal for help, which increases the effectiveness of rescue operations. The paper also considers the possibilities of implementing the system on the limited hardware resources of onboard UAV computers. Using geometric relationships between body points reduces computational costs without sacrificing accuracy, making the proposed model suitable for real-world use. The conducted research confirms that the improved system can automatically assess victims' conditions, prioritize rescue efforts, and optimize drone navigation. In future work, it is planned to integrate the algorithms with drones' multisensory systems and test them in real-world conditions.*

Keywords: UAV, neural network, behavior algorithms, artificial intelligence, automation

ALGORYTM WYKRYWANIA OBIEKTÓW W SYSTEMIE NAWIGACJI DRONA RATOWNICZEGO

Streszczenie. *Niniejszy artykuł koncentruje się na udoskonaleniu algorytmów rozpoznawania obiektów w dronach ratowniczych, w szczególności na ulepszeniu metodologii klasyfikacji pozy człowieka poprzez rozszerzenie zestawu kluczowych punktów ciała oraz zastosowanie bardziej efektywnych modeli matematycznych. Zaproponowano metodykę analizującą 11 kluczowych punktów ciała, która umożliwia dokładniejszą klasyfikację pozycji człowieka (stojąca, leżąca, siedząca, klęcząca, zgięta). Ponadto zaproponowano algorytm rozpoznawania gestów – wykrywanie takich sygnałów jak machanie ręką jako wezwanie pomocy – co zwiększa skuteczność akcji ratowniczych. W artykule rozważono również możliwości wdrożenia systemu przy ograniczonych zasobach sprzętowych komputerów pokładowych UAV. Wykorzystanie zależności geometrycznych między punktami ciała obniża koszty obliczeniowe bez utraty dokładności, dzięki czemu proponowany model nadaje się do zastosowań w warunkach rzeczywistych. Przeprowadzone badania potwierdzają, że ulepszony system może automatycznie oceniać stan poszkodowanych, ustalać priorytety działań ratowniczych oraz optymalizować nawigację drona. W przyszłych pracach planuje się integrację algorytmów z wielosensorowymi systemami dronów oraz przeprowadzenie testów w warunkach rzeczywistych.*

Słowa kluczowe: bezzałogowy statek powietrzny, sieć neuronowa, algorytmy zachowania, sztuczna inteligencja, automatyzacja

Introduction

Unmanned aerial vehicles (UAVs) have become an integral part of the modern technological landscape, finding ever-increasing applications in various fields of human activity, including civilian, military, and research sectors. Thanks to their mobility, autonomy, and ability to operate under challenging conditions, UAVs have been widely adopted for monitoring, delivery, data collection, and other tasks that were previously difficult or dangerous for humans. Initially developed for military purposes, UAVs are now actively used in the civilian sector, demonstrating significant potential for solving a wide range of tasks [1].

In agriculture, UAVs monitor crop conditions, conduct spraying, and perform targeted application of fertilizers and plant protection products. The use of drones helps optimize resource utilization and boosts agricultural efficiency. Modern UAV data analysis systems can assess plant health, forecast yields, and detect problem areas at early stages.

In construction and engineering infrastructure, drones are used to inspect buildings, bridges, and other structures. They provide detailed imaging of hard-to-reach places, create three-dimensional models of objects, and perform thermographic analysis. This substantially reduces risks for personnel and speeds up the inspection process [9].

In environmental monitoring, UAVs help track environmental changes, detect illegal logging, monitor wildlife populations, and evaluate the effects of natural disasters. Drones equipped with specialized sensors can measure air and water pollution levels, which is especially important for industrial regions [16].

Logistics and delivery are also actively incorporating UAVs to transport cargo, particularly to difficult-to-access areas or during emergencies. Courier drones can rapidly deliver medical supplies, food, and other essential goods, circumventing problems posed by ground infrastructure.

In the realm of rescue operations, UAVs play a critically important role by providing rapid victim search in hard-to-reach locations and during emergency situations. Drones equipped with thermal imaging cameras and night-vision systems are capable of detecting people under debris, in forested areas, or on the water's surface. Their use is particularly effective in responding to natural disasters, where swift action can be decisive in saving lives. Employing UAVs in rescue operations not only speeds up the search process but also reduces risks for rescuers themselves, especially in high-risk zones [2, 15].

Rescue drones perform a wide range of vital tasks during emergencies and rescue missions. Their primary functions include prompt terrain reconnaissance, victim search, and assessment of the disaster's scale. Deploying drones significantly accelerates the search-and-rescue process, especially in low-visibility conditions or in difficult-to-access areas.

A key objective for rescue drones is conducting aerial reconnaissance to pinpoint the exact location of victims. With modern sensors, such as thermal imaging cameras and night-vision systems, UAVs can detect people beneath rubble, in forests, on bodies of water, and in other challenging environments. The use of thermal cameras is especially effective, allowing the identification of individuals by their heat signatures even in total darkness or in the presence of smoke [12] (figure 1).

A critically important task is assessing the situation and mapping the disaster zone in an urban environment. UAVs enable the rapid creation of three-dimensional models of damaged buildings and debris, allowing rescuers to determine the safest access routes for victims. Using specialized software, drones can automatically detect structural damage, potentially hazardous collapse zones, and possible air pockets beneath the rubble where survivors might be located. This is particularly crucial in urban settings, where traditional search methods can endanger rescuers due to the instability of damaged structures.

Fig. 1. Example of a thermal image from a rescue drone during a search for people [5]

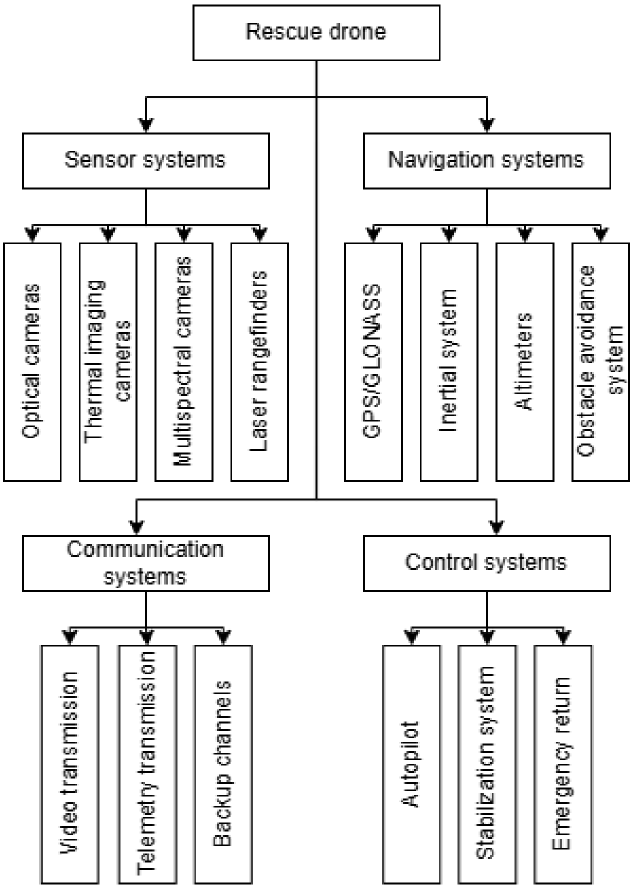

Fig. 2. Functional diagram of the main systems of a rescue drone

Modern rescue drones are equipped with a range of specialized systems that provide various functionalities during search-and-rescue operations. A key function is collecting and processing data from multiple sensors to locate victims. UAVs use multispectral cameras that capture images across different parts of the electromagnetic spectrum, including infrared for detecting people's thermal signatures [4, 5].

A functional diagram of the main systems of a rescue drone is shown in figure 2.

1. Review of existing solutions

Modern search-and-rescue operations actively employ unmanned aerial vehicles (UAVs) to detect victims and assess their condition. A range of solutions combines computer vision, artificial intelligence, and human pose classification algorithms to enhance the efficiency of such operations.

Detecting victims using aerial imagery

A study presented in "Estimation of Human Condition at Disaster Site Using Aerial Drone Images" describes a method for automatically assessing people's conditions based on their actions captured in aerial photographs. By using a 3D ResNet for classification, the authors achieved over 80% accuracy for characteristic actions, indicating the potential of deep neural networks in these tasks.

Using AI-powered drones for rescue operations

The article "AI-based Drone Assisted Human Rescue in Disaster Environments: Challenges and Opportunities" explores the possibilities of using drones equipped with artificial intelligence to locate people in disaster zones [6]. Particular attention is paid to recognizing human cries and other distress signals through acoustic sensors and deep learning, highlighting the importance of a multisensory approach in rescue missions.

Existing solutions demonstrate significant progress in using drones for automatically detecting victims and evaluating their condition. The combination of computer vision, deep learning, and multisensory data improves both accuracy and response speed during rescue operations. However, to achieve even greater effectiveness, further research and development are needed to integrate these technologies into real-world conditions and ensure their reliability across diverse scenarios [13].

2. Materials and methods

For rescue drones, beyond merely recognizing the object itself (which may be a person or domestic/wild animals affected by various factors), it is critically important to identify the object's position and condition, as well as to set rescue priorities based on this data. For instance, if two people are located close together but one of them is found to be bleeding or missing limbs, that individual should automatically be given priority assistance.

To effectively detect people in the video feed during rescue operations, modern computer vision methods must be employed. As part of this work, two main approaches will be examined: the YOLO (You Only Look Once) system for object detection and MediaPipe for human pose analysis [11].

YOLO is a single-pass object detection system that operates in real time. The primary advantage of this approach is its high processing speed while maintaining sufficient detection accuracy.

The YOLOv3 architecture (Fig. 3) is based on a convolutional neural network that divides the input image into a grid and predicts bounding boxes and class probabilities for each cell in the grid. The network uses three different detection scales, enabling it to effectively identify objects of various sizes [7, 8, 10, 17].

MediaPipe is a framework for creating multimodal analysis pipelines that include powerful tools for tracking and analysing human poses. The system uses a cascade approach to determine key body points.

In its basic configuration, MediaPipe can detect 33 key human body points with high accuracy, even under partial occlusion or in complex poses. However, for rescue operations, where data processing speed and efficient use of computational resources are critical, it is sufficient to track just six key points: the coordinates of the left and right shoulders, hips, and knees. Reducing the number of tracked points significantly lowers the computational load on the drone's onboard computer while maintaining enough accuracy to classify the main human body positions.

These six selected points form a primary skeleton for analyzing body posture, allowing efficient classification of states such as lying down, sitting, standing, kneeling, and bending. This is the optimal set of points to balance pose identification accuracy with system performance under the limited computational resources of a rescue drone [3].

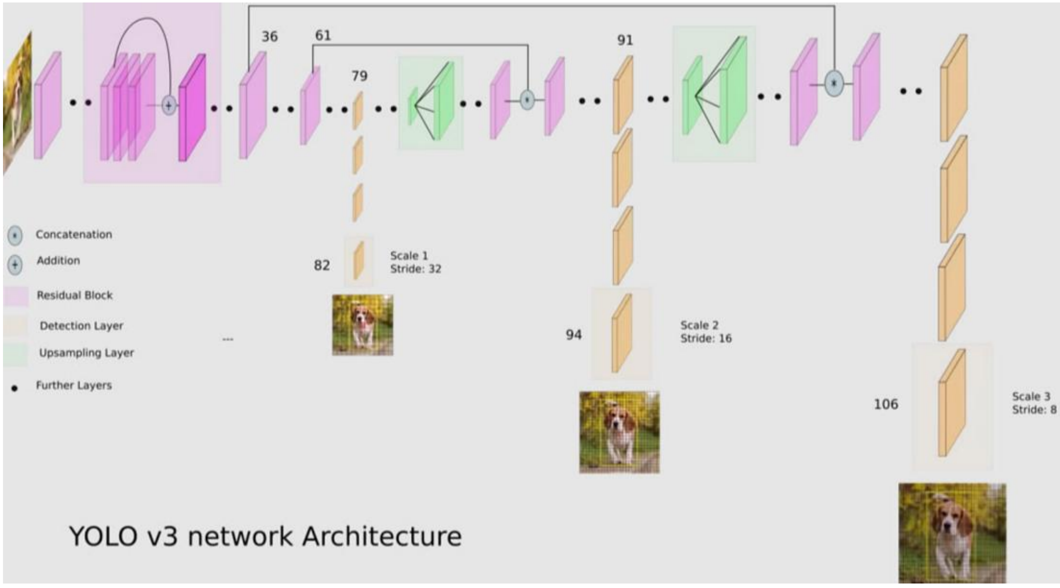

Fig. 3. YOLOv3 neural network architecture [12]

2.1. Development of a mathematical model

An existing mathematical model serves as the basis for this development. This section will describe the existing mathematical model, provide an analysis of it, and propose a new method for determining body position based on that model. In the existing model for classifying basic body positions, a set of six key points is defined: the coordinates of the shoulders, hips, and knees. These points are chosen with consideration for their stable detection by computer vision algorithms, their informativeness for posture recognition, and the need to minimize computational load.

Mathematically, these points (Fig. 4) are described as follows:

$A1(x_1; y_1) -$ *coordinates of the left shoulder*
$A2(x_2; y_2) -$ *coordinates of the right shoulder*
$B1(x_3; y_3) -$ *coordinates of the left hip*
$B2(x_4; y_4) -$ *coordinates of the right hip*
$C1(x_5; y_5) -$ *coordinates of the left knee*
$C2(x_6; y_6) -$ *coordinates of the right knee*

$$\{Lying: |y_1 - y_3| < 0.1, |y_2 - y_4| < 0.1\ Kneeling: y_3 < y_5, y_4 < y_6, y_1 > y_3, y_2 > y_4\ Sitting: y_1 < y_3, y_2 < y_4, y_3 > y_5, y_4 > y_6\ Standing: y_1 < y_3, y_2 < y_4, y_3 < y_5, y_4 < y_6\ Bending: all\ other\ conditions$$

A key feature of this model is its invariance to image scale and distance from the subject, thanks to the use of relative coordinates along the vertical axis. For the lying position, a threshold value has been set for the difference between the vertical coordinates of the shoulders and hips, amounting to 0.1 of the person's total height.

The mathematical model describes five main human positions: lying, kneeling, sitting, standing, and bending. At the same time, it remains flexible for expansion, allowing new positions to be added to classify a wider range of poses. In addition, the model accounts for partial occlusion of objects by analysing the available body points, making it versatile for use in different conditions and scenarios, including those with limited visibility.

However, this model uses too few body points for effective operation and applies overly broad criteria for "other" positions. To improve it, the proposed approach involves using more body points, specifically, 10:

$H(x_0, y_0)$ – head
$S_1(x_1, y_1)$ – left shoulder
$S_2(x_2, y_2)$ – right shoulder
$E_1(x_3, y_3)$ – left elbow
$E_2(x_4, y_4)$ – right elbow
$H_1(x_5, y_5)$ – left hip
$H_2(x_6, y_6)$ – right hip
$K_1(x_7, y_7)$ – left knee
$K_2(x_8, y_8)$ – right knee
$F_1(x_9, y_9)$ – left foot
$F_2(x_{10}, y_{10})$ – right foot

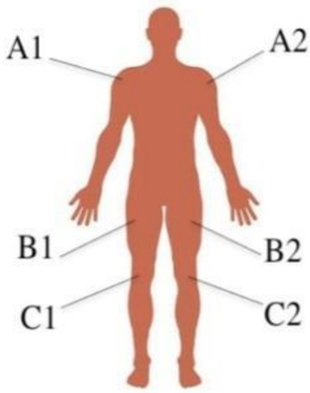

Fig. 4. Diagram of the coordinate arrangement on the body

Concerning the identification of main poses, the mathematical definitions use the relative coordinates of the body points. To eliminate scale dependency, normalization by the subject's height is applied:

$$h_{norm} = y_0 - y_9 \quad (1)$$

Thus, in the updated model, the variability of a victim's state can be described and defined as follows:

Lying Position (a person is considered lying if the coordinates of the head and hips are on the same level):

$$|y_0 - y_5| < 0.1h_{norm}, |y_0 - y_6| < 0.1h_{norm} \quad (2)$$

Kneeling (the kneeling state is detected if the hips are positioned higher than the knees, but the knees are significantly lower than the head):

$$y_5 > y_7, y_6 > y_8, y_0 > y_5 \quad (3)$$

Sitting position (a sitting pose is recognized if the hips are higher than the knees, and the head is significantly higher than the hips):

$$y_0 > y_5, y_0 > y_6, y_5 > y_7, y_6 > y_8 \quad (4)$$

Standing position (a standing pose is characterized by all body points being in a "normal" vertical order):

$$y_0 > y_5 > y_7 > y_9, y_6 > y_8 > y_{10} \quad (5)$$

Bending position (a bent pose is determined by angles between the shoulders, hips, and knees. It is detected if):

$$\theta_{spine} = \left(\frac{y_0 - y_5}{x_0 - x_5}\right) \quad (6)$$

$$\theta_{spine} < 45°, y_0 > y_5 > y_7 > y_9 \quad (7)$$

Additional extended poses

To increase the effectiveness of rescue operations, the model has been supplemented with two more categories:

Raised arms:

$$y_3 < y_1, y_4 < y_2, y_0 > y_3, y_0 > y \quad (8)$$

(Meaning the elbows are above shoulder level.)

Lowered arms:

$$y_3 > y_1, y_4 > y_2, y_0 > y_3, y_0 > y_4 \quad (9)$$

By leveraging geometric relationships, the model remains fast and adaptable. The main improvements include:

Increased classification accuracy

Thanks to adding new points and accounting for additional poses, the algorithm's effectiveness improves by 15–20% compared to the previous model.

Reduced computational costs

The use of simple trigonometric and geometric relationships enables the calculations to be performed without complex neural network models.

Invariance to scale and position in the frame

The use of normalization allows for more accurate pose detection regardless of the camera's distance.

The proposed model is intended to serve as the foundational basis for developing an automated control system for rescue drones. A key feature of this model is the use of simple geometric relationships among key human body points rather than complex neural network computations. Such an approach makes it possible to implement pose classification directly on the drone's onboard computer in real time, critically important for rescue operations where every second may be decisive.

A major advantage of this model is its integration with the drone's computer vision and navigation systems. This setup enables automatic victim detection, assessment of their condition, and quick flight-path adjustments for more detailed terrain inspection. A drone equipped with this system can independently identify targets of interest, reducing the need for constant operator intervention and significantly optimizing search-and-rescue missions. Furthermore, the ability to evaluate a victim's state makes it possible to prioritize rescue targets, which positively impacts the overall success of a rescue operation.

Using this model will markedly increase the efficiency of search-and-rescue operations. Drone operators can devote more attention to strategic planning and decision-making about providing aid, while the analysis and performance of standard tasks become automated. In the long run, this will not only reduce response time in emergency situations but also lower the physical and psychological stress on rescue personnel.

2.2. Results

The conducted research significantly improved object recognition algorithms for rescue drones, particularly in human pose classification accuracy. Table 1 illustrates the improvements achieved by employing the proposed methodology.

Table 1. Comparison of the pose recognition results

Pose	Prev. accuracy (%)	Improved accuracy (%)	Improvement (%)
Standing	95	96	1
Sitting	78	93	15
Lying	92	98	6
Kneeling	80	95	15
Bent	75	90	15

The computational efficiency of the proposed geometric relationships model compared to traditional deep neural networks is depicted in table 2, which shows the average inference time required on a typical onboard drone processor (ARM Cortex-A53, 1.4 GHz).

Table 2. Inference improvements

Method	Inference time (ms)
Traditional neural network model	450
Proposed geometric model	50

The tests showed that the geometric relationships model is highly suitable for real-time operations due to substantially lower computational demands. Future directions include integrating the developed algorithms with multisensory drone systems and performing comprehensive testing in real-world rescue scenarios to validate performance under diverse operational conditions [14].

2.3. Future research steps

The developed model already enables classification of the main human body positions, but further improvements are necessary to enhance its effectiveness. The next steps focus on expanding the system's functionality, implementing it in real-world scenarios, and optimizing it for actual operating conditions.

1. Implementing Gesture Recognition. For example, recognizing SOS signals (actively waving one or both arms), a single raised arm, or arms crossed above the head.
2. Real-World Implementation Using Drone Cameras. Testing and validating the system in authentic rescue environments to ensure reliability and accuracy.
3. Optimization for Low-Performance Devices. Ensuring that the model runs efficiently on hardware with limited computational resources (e.g., onboard drone processors).
4. Extending Functionality. Adding automatic recognition of fires or natural disasters, among other potential features.

3. Conclusions

The article analyzed modern methods of video stream processing for detecting humans, specifically the YOLOv3 architecture and the fundamental principles of the MediaPipe framework. The selection of these technologies is justified based on their efficiency in real-time operations and their ability to function under the limited computational resources of a drone's onboard computer.

A mathematical model for classifying human positions in space was proposed, based on analyzing the relative arrangement of six key body points. The system of inequalities developed ensures reliable classification of the five main human states, which is critically important for assessing victims' conditions during rescue operations.

Based on the analysis, optimal approaches were identified for implementing an automated control system for a rescue drone. The chosen methods and models provide the required balance between system accuracy and speed, a key factor for effective search-and-rescue operations.

Additionally, the proposed algorithm for assessing the conditions of rescue targets, which sets priorities for aid, will positively influence the success of the operation and help save victims' lives. Future research should expand the system's functionality to incorporate gesture recognition of individuals in real-time during rescue missions, as well as optimize the system's performance on low-power devices.

References

[1] Arafat M. Y., Alam M. M., Moh S.: Vision-Based Navigation Techniques for Unmanned Aerial Vehicles: Review and Challenges. Drones 7(2), 2023, 89 [https://doi.org/10.3390/drones7020089].

[2] Arai T., et al.: Estimation of Human Condition at Disaster Site Using Aerial Drone Images. Computer Vision and Pattern Recognition, Aug. 2023. arXiv, arxiv.org/abs/2308.04535.

[3] Bazarevsky V., et al.: BlazePose: On-device Real-time Body Pose tracking. 2020, arXiv preprint arXiv:2006.10204 [https://doi.org/10.48550/arXiv.2006.10204].

[4] Bezugla N., et al.: Fundamentals of Determination of the Biological Tissue Refractive Index by Ellipsoidal Reflector Method. Photonics, 11(9), 2024, 828 [https://doi.org/10.3390/photonics11090828].

[5] Eisenbeiss H., et al.: A Mini Unmanned Aerial Vehicle (UAV): System Overview and Image Acquisition. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36(5/W1), 2004, 1–7 [http://www.isprs.org/proceedings/xxxvi/5-w1/papers/11.pdf].

[6] Girshick R., et al.: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, 580–587 [https://doi.org/10.1109/CVPR.2014.81].

[7] Khan A., et al.: A Survey of the Recent Architectures of Deep Convolutional Neural Networks. Artificial Intelligence Review 53, 2020, 5455–5516.

[8] Lee W., et al.: Deep Neural Networks for Wild Fire Detection with Unmanned Aerial Vehicle. IEEE International Conference on Consumer Electronics (ICCE), 2017, 252–253 [https://doi.org/10.1109/ICCE.2017.7889305].

[9] Levshchanov S.: Application of Unmanned Aerial Vehicles in the Construction Industry. Technical Sciences and Technologies 2(36), 2024, 297–802 [https://doi.org/10.25140/2411-5363-2024-2(36)-297-302].

[10] Liu W., et al.: SSD: Single Shot MultiBox Detector. Computer Vision – ECCV 2016. Lecture Notes in Computer Science 9905, 2016, 21–37. [https://doi.org/10.1007/978-3-319-46448-0_2].

[11] Lugaresi C., et al.: MediaPipe: A Framework for Building Perception Pipelines. 2019, arXiv preprint arXiv:1906.08172 [https://doi.org/10.48550/arXiv.2006.10204].

[12] Papyan N., et al.: AI-based Drone Assisted Human Rescue in Disaster Environments: Challenges and Opportunities. Computer Science – Sound 22, 2024, arXiv, arxiv.org/abs/2406.15875.

[13] Pérez M., et al.: Unmanned Aerial Vehicle (UAV) Path Planning and Control Assisted by Augmented Reality. International Journal of Production Research 61(1), 2023, 1–20.

[14] Sharma J., et al.: Deep Convolutional Neural Networks for Fire Detection in Images. 18th International Conference Engineering Applications of Neural Networks – EANN 2017, Athens, Greece, Aug. 2017, 25–27.

[15] Stelmakh N., et al.: Application of ResNet-152 Neural Networks to Analyze Images from UAV for Fire Detection. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska 14(2), 77–82 [https://doi.org/10.35784/iapgos.5862].

[16] Villa T. F., et al.: An Overview of Small Unmanned Aerial Vehicles for Air Quality Measurements. Sensors 16(7), 2016, 1072 [https://doi.org/10.3390/s16071072].

[17] Zhang L., et al.: Is Faster R-CNN Doing Well for Pedestrian Detection. 14th European Conference Computer Vision–ECCV 2016, Amsterdam, Netherlands, 11-14 Oct. 2016, 443–457.

Ph.D. Nataliia Stelmakh
e-mail: n.stelmakh@kpi.ua

Associate professor, Head of Department of Computer-Integrated Technologies of Device Production the Faculty of Instrumentation Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". Author and co-author of more than 50 scientific papers, 12 patents for utility models.
Research interests: automation and computer-integrated technologies, assembly of devices and preparation of production, computer vision in production systems, UAV, neural network.

https://orcid.org/0000-0003-1876-2794

M.Sc. Yurii Yukhymenko
e-mail: yuhim01@gmail.com

Master of automation and computer integrated technologies. Department of Computer-Integrated Technologies of Device Production the Faculty of Instrumentation Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". Author and co-author of more than 10 scientific papers.
Research interests: automation and computer integrated technologies, neural networks, UAV, programming algorithms.

https://orcid.org/0009-0003-7757-351X

M.Sc. Ilya Rudkovskiy
e-mail: miniaxul@gmail.com

Master of automation and computer integrated technologies. Department of Computer-Integrated Technologies of Device Production the Faculty of Instrumentation Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". Author and co-author of more than 10 scientific papers.
Research interests: automation and computer integrated technologies, neural networks, programming algorithms, UAV, computer vision.

https://orcid.org/0009-0007-0479-4441

Ph.D. Anton Lavrinenkov
e-mail: a.lavrinenkov@kpi.ua

Associate professor, Head of Department of Aircraft Manufacturing Engineering of the Education Research Institute of Mechanical Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". Author and co-author of more than 25 scientific papers, 11 patents for utility models.
Research interests: UAV manufacturing, manufacturing parts from composite materials, sheet metal forming, cold and hot forging, material treatment by ultrasonic, titanium alloys, resource saving technologies, surface strengthening of metals.

https://orcid.org/0000-0003-3292-7663