Information, Computing and Intelligent Systems, 2025, No. 7, 83 – 95
UDC 004.8 https://doi.org/10.20535/2786-8729.7.2025.341779

UAeroNet: DOMAIN-SPECIFIC DATASET FOR AUTOMATION OF UNMANNED AERIAL VEHICLES

Yuriy Kochura*
https://orcid.org/0000-0002-4217-8152

Yevhenii Trochun
https://orcid.org/0000-0002-2744-6681

Vladyslav Taran
https://orcid.org/0000-0003-2493-7239

Yuri Gordienko
https://orcid.org/0000-0003-2682-4668

Oleksandr Rokovyi
https://orcid.org/0000-0001-6934-7502

Sergii Stirenko
https://orcid.org/0000-0001-5478-0450

National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

*Corresponding author: kochura@comsys.kpi.ua

This paper addresses the challenges and key principles of designing domain-specific datasets that can be used especially for automation of unmanned aerial vehicles. Such datasets play a key role in building intelligent systems that enable autonomous operation and support data-driven decisions. The study presents approaches we used for data collection, analysis and annotation, highlighting their importance and practical impact on real-world application. The preparation of a domain-specific dataset for automating unmanned aerial vehicles operations (such as navigation and environmental monitoring) is a challenging task due to frequently low image resolution, complex weather conditions, a wide range of object scales, background noise and heterogeneous terrain landscapes. Existing open datasets typically cover only a limited variety of unmanned aerial vehicles use cases, which restricts the ability of deep learning models to perform adequately under non-standard or unpredictable conditions.

The object of the study is video data acquired by unmanned aerial vehicles for creating domain-specific datasets that enable machine learning models to perform autonomous object recognition, navigation, obstacle avoidance and interaction with an environment with minimal operator involvement. The subject focuses on the collection, preparation and annotation of video data acquired by unmanned aerial vehicles. The purpose of the study is to develop and systematize workflow for creating specialized datasets to train robust models capable of autonomously recognizing objects in real-time video captured by unmanned aerial vehicles. To achieve this goal, a workflow was designed for collecting and annotating video data, raw video data were acquired from unmanned aerial vehicles sensors and manually annotated using the Computer Vision Annotation Tool.

As a result of this work, we developed a domain-specific dataset (`UAeroNet`) using an open-source annotation tool for object tracking task in real scenarios. `UAeroNet` consists of 456 annotated tracks and a total of 131 525 labeled instances that belong to 13 distinct classes.

Keywords: unmanned aerial vehicles, UAeroNet, object detection, autonomous navigation, computer vision.

1. Introduction

Unmanned aerial vehicles (UAVs) are increasingly being used to collect data in the form of video streams, which is of great importance in many areas, such as agriculture, environmental monitoring, territory protection and rescue operations. Effective use of this data requires the implementation of a system that automates the processes of collecting, processing and annotating images obtained from UAV sensors like visible light cameras and infrared (IR) video cameras.

A large amount of reliable and accurate data is required to create automated methods of UAV operation based on artificial intelligence and edge computing [1–4]. Such data can ensure effective training of deep neural networks and guarantee high accuracy of model predictions when UAVs perform navigation, environmental monitoring, rescue operations [5, 6], etc.

The relevance of this topic is that UAVs generate massive amounts of visual data, and developing automated systems for processing and analysing this data is essential for improving operational efficiency, enabling real-time decision-making, which is critical for a wide range of applications.

In conclusion, research on automated processing of UAV-acquired data is highly relevant, as it enhances the accuracy and speed of decision-making across various fields and provides a foundation for the further development of intelligent autonomous UAV systems.

2. Literature review and problem statement

The preparation of domain-specific datasets, particularly for UAV applications, where computational resources are limited [7], is a challenging task due to factors such as low image resolution, complex weather conditions, variable object scales, environmental noise and diverse terrains. The data volume of the prepared dataset is usually the key to reaching better model accuracy and generalisation.

We have identified common problems that most often arise when preparing a high-quality dataset for autonomous UAV operations:

1. Traditional machine learning models lack contextual understanding. Understanding the context would allow the UAV to take into account the specifics of each situation and extrapolate the knowledge gained to make the right decisions. Thus, for some tasks, it may be necessary to take into account the context of the scene during data preparation (see Fig. 1).

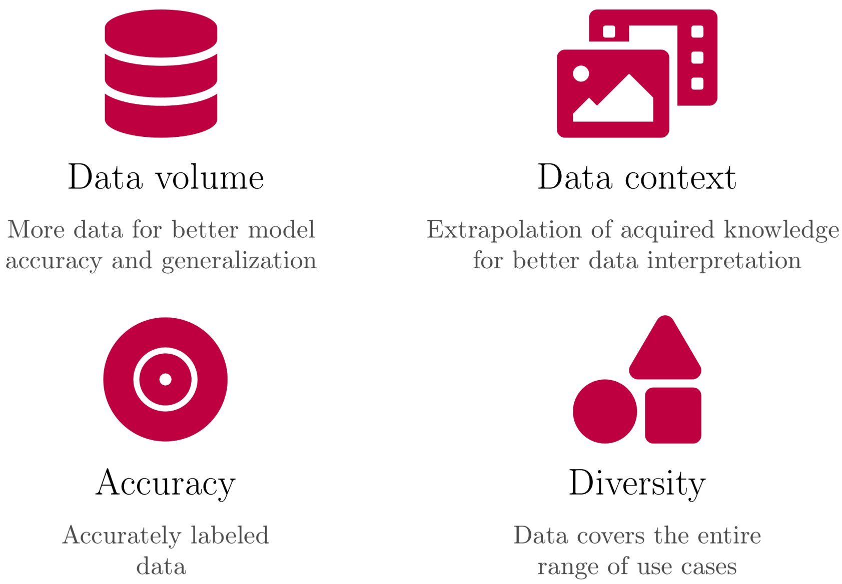

Fig. 1. Prerequisites for creating a high-quality dataset

2. Models derived from supervised learning can't self-correct true label errors that may have been made during the data preparation stage, so once the model learns the template for object recognition, this knowledge (whether it is wrong or not) remains embedded in the model parameters. Therefore, the more accurately and clearly the data is annotated, the better the model's predictions will be, which will enable the UAV to perform its mission more efficiently and provide a higher level of autonomy and safety.

3. The limited variety of available data is a key challenge in creating a comprehensive, specialized dataset for UAV use. Existing datasets often do not cover all possible use cases, which limits the

ability of models to perform adequately in non-standard or unpredictable conditions. This poses a challenge in ensuring the versatility and robustness of machine learning algorithms in real-world field conditions [8–10].

The reverse engineering of the UAV-based object perception model may allow to determine which data preparation errors are acceptable and which are not, as well as to identify best practices for annotating data obtained from UAV sensors.

The unresolved problem is the lack of freely available, comprehensive, context-aware and accurately annotated datasets for training robust machine learning models, which may be used in UAV-specific scenarios. Traditional approaches lack contextual understanding of the scene and it's difficult to correct label errors introduced during annotation, which reduces prediction reliability. Furthermore, the limited diversity of available datasets restricts model adaptability to real-world, non-standard conditions, including variations in the technical characteristics of the sensors used in the UAV, the altitude and speed of the UAV, the type of terrain, lighting conditions, time of year, weather factors such as wind, snow and rain as well as the presence of various types of obstacles.

3. The aim and objectives of the study

The aim of this study is to develop and systematize a workflow for creating and processing specialized datasets to enable the training of models capable of autonomously recognizing objects in real-time video captured by UAVs. The usage of such models will help to increase the level of autonomy and efficiency of UAV tasks by integrating machine learning algorithms into UAVs that will identify, track and classify objects of interest during the mission. This is especially important for autonomous systems that need to follow a target, avoid obstacles and interact with the environment without or with minimal operator involvement.

To achieve this goal, the following objectives were defined:

- design a workflow for annotating and processing UAV-acquired video data;
- identify relevant data sources;
- acquire raw video data from UAV sensors;
- perform manual annotation of the data using Computer Vision Annotation Tool (CVAT) for object detection and tracking tasks.

4. The study materials and methods of the development of a domain-specific dataset

A high-performance data annotation pipeline requires a strategic combination of labor force with the technical knowledge, tools and operations that can ensure the high quality of the data produced. Fig. 2 summarizes the key elements of a data annotation organization process. Effective coordination between these components enables consistent, accurate and scalable annotation outcomes.

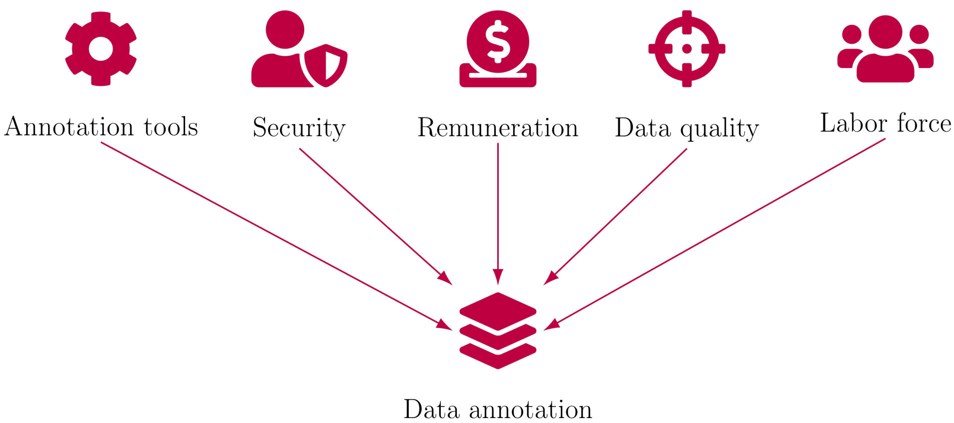

Fig. 2. Key elements of the data annotation process

4.1. Annotation tool

An annotation tool (specialized software) is a necessary component of any data labeling process. There are specialized tools (open source or commercial software) for different types of annotation, so it becomes obvious that the choice of a particular tool will depend on its usability and compliance with the standard requirements of the task to be solved.

The choice of an annotation tool should be guided by several important considerations. First, available tools must be evaluated for their relevance to the specific annotation task. Another key decision concerns whether to develop a custom solution, purchase a commercial product or use an existing open-source software. Developing a tool from scratch or adapting an existing one offers greater flexibility for managing changes, customizing it to the task's specific needs and ensuring maximum control and data security. In addition, effective annotation tools typically offer seamless integration with modern technologies, allowing import and export of annotations in widely used formats such as You Only Look Once (YOLO), Common Objects in Context (COCO), PASCAL Visual Object Classes (VOC), as well as compatibility with popular machine learning and computer vision frameworks including `TensorFlow`, `PyTorch` and `Open Source Computer Vision Library (OpenCV)`. Finally, an intuitive and well-designed interface remains essential to ensure efficient and accurate annotation.

4.2. Security

Data security during annotation is an important aspect, especially when it comes to sensitive, confidential or personal information. The data annotation service must meet regulatory and other requirements depending on the level of security the data requires. There should be a facility where the work can be done safely, with proper employee training, processes, policies and procedures to follow during data annotation. Proven and secure platforms that maintain security and quality standards should be used for data annotation. Additionally, data must be encrypted during both transmission and storage.

Data security during the annotation process can be reinforced through a combination of organizational and technical measures. Annotators are typically required to sign non-disclosure agreements or equivalent documents specifying data handling obligations. In addition, providing training on data security protocols and monitoring compliance helps maintain adherence to security standards. Technical safeguards, such as disabling write access on annotation devices, further reduce the risk of data leakage. For on-site operations, physical security measures, including video surveillance of the annotation workspace, provide an additional layer of protection.

4.3. Remuneration

Another important component of data annotation is labor pricing. The model that is used to determine the price of labor can affect the overall cost and quality of the data produced. Pricing is a complex process that must take into account the type and quality of data to be annotated, the type of annotations to be performed (segmentation, bounding boxes, labels, etc.), the number of classes, the timeframe and the amount of data. More data usually means more cost. This can be the number of images, videos or other information. If the data requires detailed annotation, it will affect the overall cost of the project. Different types of annotation have different levels of complexity and therefore different prices. Projects with high requirements for the accuracy of the prepared data may include quality checks, which will increase the overall cost of data preparation. More experienced labelers can perform tasks more accurately, but their services may be more expensive. Some projects may require specialized knowledge to prepare accurate data, which will affect the cost of data annotation services.

In practice, the most commonly employed pricing models are based either on the volume of annotated units (e.g., objects, frames) or on an hourly rate, which is preferred when data volume or annotation complexity is uncertain.

4.4. Data quality

The quality of labels is the most important part of the data annotation process, affecting the quality of the entire prepared dataset. Optimizing the data annotation process to obtain the highest quality with the available resources is a continuous process as the quality of the annotation depends on various factors: the functionality of the tool used to annotate the data, ambiguity in the data labeling instructions, the experience of the labelers, the workflow for ensuring the quality of data preparation, and the type and nature of the data to be labeled.

Data annotation guidelines are extremely important to ensure the proper quality of the data produced. They serve as a guide for labelers, helping them to correctly interpret and process data in accordance with project requirements. Instructions help ensure that all labelers follow the same rules and principles, which is especially important when working with a large team. The use of uniform guidelines helps to reduce the number of errors and achieve a unified approach to annotation, regardless of who is doing the work.

4.5. Labor force

A large volume of data that need to be prepared for further training of machine learning methods require the involvement of an adequate amount of labor to meet the needs of the project. Efficiency and scalability in annotation are influenced by the relationship between workforce size and dataset volume, as well as by the allocation of personnel according to labeling frequency. Evaluating annotator performance in terms of speed and accuracy provides insights into process effectiveness, while continuous refinement based on feedback and error analysis contributes to improved annotation quality.

When selecting personnel for data annotation, institutions may rely on several approaches. Internal employees can be engaged to perform labeling tasks, providing direct control over the process. Alternatively, institutions may employ external groups of annotators, such as cloud workers, or contract specialized third-party data labeling companies. Another option involves leveraging online crowdsourcing platforms to access distributed annotators for large-scale labeling tasks.

Thus, the first step is to recruit people with the appropriate skills and knowledge to annotate the data. These can be either internal employees or freelancers. New annotators are trained on established rules and standards with illustrative examples of correct and incorrect annotations to ensure clarity. Consistent processes and standards are critical for maintaining annotation quality and continuous improvement. Quality assurance procedures, including repeated verification of labeled data, support the provision of constructive feedback to annotators, fostering higher accuracy in subsequent work. Coordination of the labeling team, task assignment, and monitoring of task completion further contribute to the efficiency and reliability of the annotation process.

4.6. Active learning strategy

The most time-consuming, but the most reliable approach to annotating data to improve model accuracy is to manually label all available data (creating ground truth labels for the machine learning model). Data annotation efficiency can be enhanced through the use of specialized annotation tools with intuitive interfaces that minimize labeler workload, regular audits to maintain labeling accuracy and the integration of the active learning method [11] that identify the most informative samples from unlabeled data.

Active learning is a machine learning method in which the model has the ability to choose which data to use for training. This can significantly reduce the cost and volume of data that needs to be annotated manually, while maintaining high model accuracy. The main idea of active learning is for the model to identify examples from the dataset with the least number of labels represented in the dataset, or examples for which the model is least confident in its predictions (see Fig. 3). These examples should be annotated manually to improve the overall performance of the model. The active learning algorithm should be repeated starting from step 3 (see Fig. 3) until the model reaches the required accuracy.

Step 1
Annotate the raw data

Step 2
Train the model on prepared data

Step 3
Identify samples with a low level of confidence

Step 4
Annotate each sample manually and add it to the training set

Step 5
Retrain the model with new data

Step 6
Control annotation process and model accuracy

Fig. 3. Active learning algorithm [11]

This approach is especially useful in cases where data annotation is an expensive and time-consuming process. However, active learning requires an initial set of labeled examples to initiate the selection process. If this initial dataset is unrepresentative or biased the model may iteratively select suboptimal samples.

5. Results of the development of a domain-specific dataset for UAVs automation

5.1. Workflow for data acquisition, annotation and preprocessing

The workflow (Fig. 4) illustrates the sequential process of acquiring, labeling and preprocessing data for UAV automation tasks. It begins with assessing data availability and the possibility of using external or synthetic sources to ensure sufficient volume and diversity. Next, the data undergoes annotation and quality improvement. Depending on dataset size and available resources, labeling may involve semi-supervised learning, weak supervision or crowdsourcing.

The final stage involves preprocessing, including data cleaning to reduce noise, data transformation to meet model input requirements and feature engineering to improve model learnability.

5.2. Data sources and collection of UAV data

We conducted a search for, collected and analyzed visual data for further training and development of machine learning models for UAV onboard applications. The search for relevant data included open sources: groups and channels on social media platforms and cross-platform instant messaging systems (Facebook, YouTube, WhatsApp, Instagram, Telegram, etc.), the data science competition platform (Kaggle), Papers with Code, the developer platform (GitHub), etc.

Collected data were evaluated in terms of type (educational, competition, raw), data volume, representativeness of real-world scenarios and suitability for labeling (see Tab. 1). Based on this

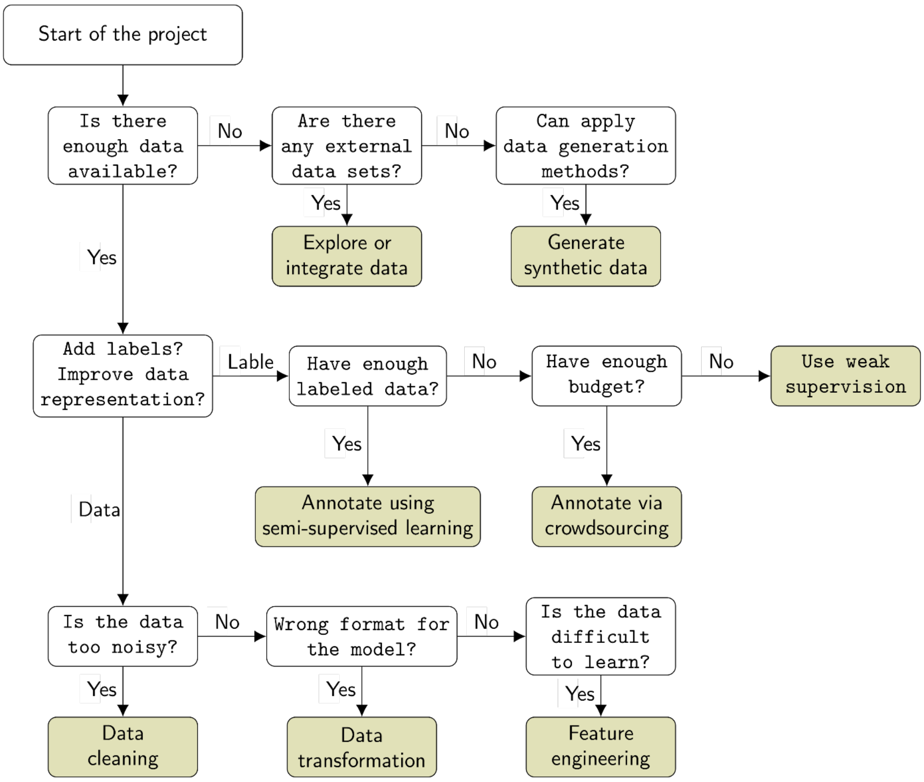

Fig. 4. Flowchart of data acquisition, annotation and preprocessing

Table 1. General characteristics of datasets: advantages and disadvantages

Dataset type	Advantages	Disadvantages
Educational (academic)	Clear data, easy to use	Limited data sample, overly simplified, results are not always transferable to real-world conditions
Competition	Data is close to real	Still simplified data and available only for hot topics, the results are not always transferable to real-world conditions
Raw data	Great flexibility	It takes a lot of effort to prepare

evaluation, raw data that meet the project requirements were chosen and annotated, while model training on these data is planned in a future study.

Below are several academic and competition datasets for civilian use of UAVs for various purposes, including research, development and testing of UAVs. Each dataset was assessed in terms of its relevance, coverage of use cases, data quality and suitability for integration into the project workflow.

5.2.1. The UAV123 dataset

The videos captured from UAVs at low altitude [12] are inherently different from videos from popular object tracking datasets such as `OTB50`, `OTB100`, `VOT2014`, `VOT2015`, `TC128`, and `ALOV300++`. Therefore, a new dataset (`UAV123`) for long-term aerial object tracking (`UAV20L`) was proposed [12].

The `UAV123` dataset contains a total of 123 video segments and more than 110 thousand frames, making it the second largest object tracking dataset after `ALOV300++`. All videos are annotated with upright bounding boxes. A few examples from the dataset are shown in the Fig. 5.

Fig. 5. Examples of images from the UAV123 dataset [12]

5.2.2. The UAVDT dataset

This dataset contains images and videos captured from UAVs in urban and rural areas, annotated for object detection and tracking tasks [13].

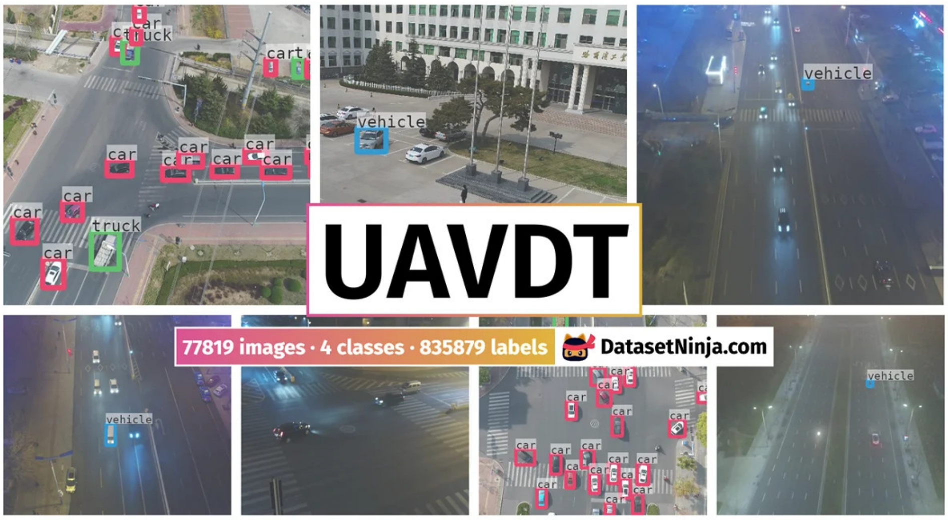

Fig. 6. Examples of images from the UAVDT dataset [13]

In total, the `UAVDT` dataset includes 77819 frames featuring 835879 annotated objects belonging to four distinct categories: car, vehicle, truck and bus. Several examples from the dataset are shown in the Fig. 6.

5.2.3. The DAC-SDC dataset

Design Automation Conference System DesignContest 2022 Dataset (`DAC-SDC`) [14] is a dataset for object detection and identification tasks. It is used in the field of search and rescue (SAR). The dataset consists of 93520 images and 93520 object labels belonging to 12 different classes, including: person, car, riding, boat, group, wakeboard, drone, truck, paraglider, whale, building, and horseride [14].

Fig. 7. Examples of images from the DAC-SDC dataset [14]

The dataset was published in 2022 by the University of Notre Dame, USA, Peking University, China and the University of Pittsburgh, USA [14]. A few examples from the dataset are shown in the Fig. 7.

5.2.4. The VisDrone2019-DET dataset

`Drone2019-DET` [15] is a dataset for object detection tasks. Some examples from the dataset are shown in the Fig. 8.

Fig. 8. Examples of images from the VisDrone2019-DET dataset [15]

The dataset consists of 10209 images and 471266 labels belonging to 13 different classes, including: car, pedestrian, van, motor, person, truck, ignored region, bicycle, bus, tricycle, awning tricycle, other and people [15].

5.2.5. Raw data

Academic and competition datasets (see Table 1) can't fully satisfy the requirements for developing robust models capable of deployment in real workflows. Although integrating multiple academic datasets could partially mitigate data limitations, differences in data preparation pipelines often hinder such integration and reduce data consistency. Consequently, harmonization and reuse of these datasets remain challenging. Another potential approach is the generation of synthetic data.

However, this method has several drawbacks, as generating synthetic data can be challenging and may not always capture the full diversity and complexity of real data. In addition, producing high-quality synthetic data can be computationally expensive, particularly for domain-specific environments. Therefore, academic and competition datasets, as well as synthetic data generation, are treated as supplementary options, since none of them can serve as a reliable basis for developing models intended for deployment in the wild.

The conditions of data acquisition for UAVs depend on several factors, including the technical characteristics of the sensors used in the UAV, the altitude and speed of the UAV, the terrain, lighting conditions, time of year, weather conditions (such as wind, snow and rain) and artificial obstacles. For applied, domain-specific AI tasks, engineers must rely on raw data collected under real conditions. Raw data represent features in their original form of acquisition, often unstructured. Using raw data for annotation ensures that the trained artificial intelligence (AI) model, once deployed, will process input data in the same format as during training. Such consistency provides a solid foundation for achieving high model adaptability and performance.

However, for this project, there are currently no publicly available datasets containing UAV-acquired video data collected under aggressive environmental conditions. Therefore, the most appropriate approach was to search for raw data obtained in challenging conditions using aerial drones operating at different altitudes, under various weather and lighting conditions, across diverse landscapes and with different types of onboard equipment. To accomplish this task, we researched open sources: groups and channels on social media platforms and cross-platform instant messaging systems (Facebook, YouTube, WhatsApp, Instagram, Telegram), a data science competition platform (kaggle), paperswithcode and a developer platform (github). As a result, more than 247 videos were collected, which were filmed from the UAV using an optical camera and more than 55 videos using an IR camera.

It is worth noting that not all collected videos are presented in good resolution, some videos contain additional attributes, captions on the frames, some of which were added by the authors of these videos during additional processing, which negatively affects the overall quality of the data and can complicate the process of label preparation.

5.3. Annotation tool used for object detection and tracking tasks

We used `CVAT` [16] to centrally store UAV sensor images, preprocess them (adjust brightness, contrast, color balance) and create datasets with manual or automated markup. `CVAT` is an open-source tool developed by Intel and distributed under the MIT License [16], which allows to freely use, modify and redistribute this software for both commercial and non-commercial purposes. `CVAT` supports multi-user mode, offers easy integration with modern technologies and provides tools for creating polylines, polygons, rectangles, and other geometric shapes necessary for accurate object labeling in computer vision applications.

The `CVAT` interface includes the following main components: the navigation and menu bar, the control panel, the workspace and the object panel (see Fig. 9).

5.4. Prepared dataset

We have performed object detection using bounding boxes in Track mode (each new object is accompanied by a bounding box until the end of the episode).

The prepared dataset consists of 52 videos, 456 tracks, 131525 annotated objects belonging to 13 unique classes. During annotation, 19 classes (objects of interest) were used.

5.5. Annotation guidelines

We followed the following rules when annotating the data:

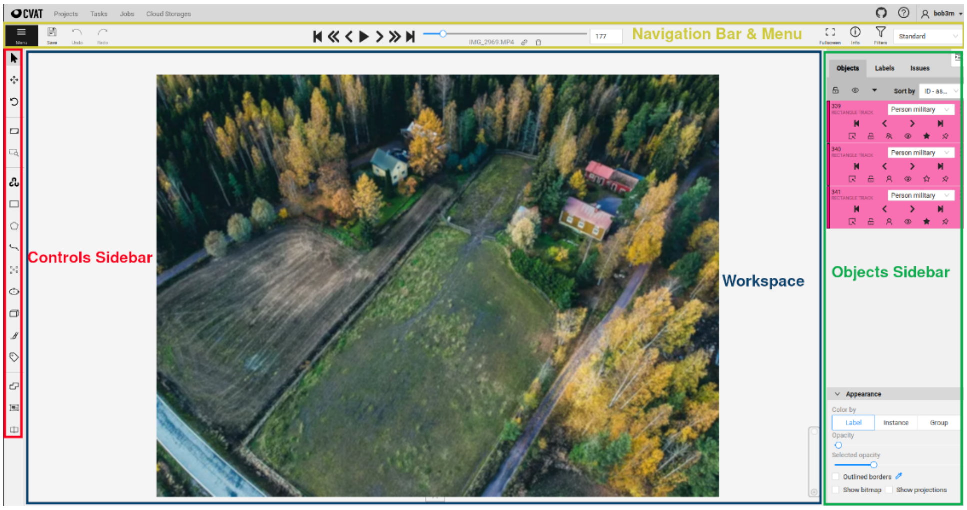

Fig. 9. CVAT interface

1. Don't create a bounding box for a completely overlapping object. If an object is completely covered by another object, do not draw a bounding box.
2. When an object is partially covered by another object or when part of the object is off-screen, turn on the occluded property.
3. If an object disappears from view and is not visible for several frames and then reappears, use the Merge function to combine several separate tracks into one.
4. The bounding box should cover the entire object of interest with the smallest possible area.
5. Remove video fragments (frames) that do not contain objects of interest: Alt + Del.

6. Discussion of results

While preparing the specialized dataset, we faced with some challenges. Some videos were of low quality due to poor weather or lighting conditions, which made annotation difficult. UAVs capture complex scenes from different heights and angles, covering urban, rural and natural environments that is why objects such as persons, buildings, vehicles and so on look different depending on the conditions of the video recording, which makes annotation difficult too. Accurate annotation of objects in highly cluttered environments (e.g., forests and cities) requires high skill and experience of the labeler. Usually, annotating small objects is very challenging due to low video resolution or object distortion. Sometimes, due to poor video quality or small size of the object, it was difficult to determine the true label. Analysis of the prepared dataset shows that some classes are much less represented compared to others. This can lead to problems when training AI models, which may overly focus on more frequent classes and fail to cope with rare classes. Further improvements will involve expanding the dataset and increasing the number of examples for underrepresented classes to partially compensate and address the imbalance in the data.

Conclusions

The study aimed to develop and systematize a workflow for creating and processing specialized datasets to enable the training of models capable of autonomously recognizing objects in real-time UAV video. This goal was successfully achieved through the design and implementation of a structured process covering data acquisition, preparation and annotation.

A workflow for annotating and processing UAV-acquired video data was designed and practically applied. Relevant data sources were identified through an extensive search of open platforms, and raw video data were collected from UAV sensors operating under diverse real-world conditions. The acquired data were manually annotated according to established rules to ensure consistency and accuracy. Annotation was performed using CVAT for 19 objects of interest. As a result, a specialized dataset (UAeroNet) was prepared, containing 52 videos, 456 object tracks and 131525 annotated objects across 19 classes of interest.

The developed workflow and dataset provide a foundation for further research aimed at training deep learning models capable of autonomous object recognition and tracking in real time. Future work will focus on expanding the dataset, balancing underrepresented classes and integrating trained models into UAV systems to enhance their autonomy and operational efficiency.

Acknowledgements

This work was supported by the Ministry of Education and Science of Ukraine under the project No. 2715r.

References

[1] Y. Gordienko, N. Gordienko, R. Fedunyshyn, and S. Stirenko. "Multibackbone ensembling of EfficientNetV2 architecture for image classification," *in International Conference on Trends in Sustainable Computing and Machine Intelligence*, pages 55–69, Springer, 2024, http://dx.doi.org/10.1007/978-981-96-1452-3_5.

[2] Z. Ouyang, J. Niu, T. Ren, Y. Li, J. Cui, and J. Wu, "Mbbnet: An edge iot computing-based traffic light detection solution for autonomous bus," *Journal of Systems Architecture*, vol. 109, p. 101835, 20205, https://doi.org/10.1016/j.sysarc.2020.101835.

[3] A. R. Khouas, M. R. Bouadjenek, H. Hacid, and S. Aryal, "Training machine learning models at the edge: A survey," *arXiv preprint*, arXiv:2403.02619, 2024.

[4] Y. Matsubara and M. Levorato, "Neural compression and filtering for edge-assisted real-time object detection in challenged networks," *in 2020 25th International Conference on Pattern Recognition (ICPR). IEEE*, 2021, pp. 2272–2279, https://doi.org/10.48448/830a-4b92.

[5] A. Polukhin, Y. Gordienko, G. Jervan, and S. Stirenko, "Object detection for rescue operations by high-altitude infrared thermal imaging collected by unmanned aerial vehicles," *in Iberian Conference on Pattern Recognition and Image Analysis. Springer*, 2023, pp. 490–504, http://dx.doi.org/10.1007/978-3-031-36616-1_39.

[6] R. Murphy and T. Manzini, "Improving drone imagery for computer vision/machine learning in wilderness search and rescue," *in 2023 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR). IEEE*, 2023, pp. 159–164, http://dx.doi.org/10.48550/arXiv.2309.01904.

[7] A. Polukhin, Y. Gordienko, M. Leier, G. Jervan, O. Rokovyi, O. Alienin, and S. Stirenko, "Edge intelligence resource consumption by uav-based ir object detection," *in Proceedings of the 2023 Workshop on UAVs in Multimedia: Capturing the World from a New Perspective*, 2023, pp. 57–61, https://doi.org/10.1145/3607834.3616566.

[8] N. Gordienko, Y. Gordienko, and S. Stirenko, "Enhancing deep learning sustainability by synchronized multi augmentation with rotations and multi-backbone architectures," *Big Data and Cognitive Computing*, vol. 9, no. 5, p. 115, 2025, https://doi.org/10.3390/bdcc9050115.

[9] K. A. Hambarde, N. Mbongo, P. K. MP, S. Mekewad, C. Fernandes, G. Silahtaroğlu, A. Nithya, P. Wasnik, M. Rashidunnabi, P. Samale et al., "Detreidx: A stress-test dataset for real-world uav-based person recognition," *arXiv preprint* arXiv:2505.04793, 2025.

[10] J. Xiao and G. Loianno, "Uasthn: Uncertainty-aware deep homography estimation for uav satellite-thermal geo-localization," *arXiv preprint* arXiv:2502.01035, 2025.

[11] D. Cohn, L. Atlas, and R. Ladner, "Improving generalization with active learning," *Machine learning*, vol. 15, no. 2, pp. 201–221, 1994, https://doi.org/10.1007/BF00993277.

[12] Image and V. U. Lab, "Uav123 & uav20l," 2024, [Online]. Available: https://cemse.kaust.edu.sa/ivul/uav123.

[13] "Visualization tools for uavdt dataset," [Online]. Available: https://datasetninja.com/uavdt.

[14] "Visualization tools for dac-sdc 2022 dataset," [Online]. Available: https://datasetninja.com/dac-sdc-2022.

[15] P. Zhu, L. Wen, D. Du, X. Bian, H. Fan, Q. Hu, and H. Ling, "Detection and tracking meet drones challenge," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 7380–7399, 2021, http://dx.doi.org/10.1109/TPAMI.2021.3119563.

[16] CVAT.ai Team. Computer vision annotation tool (cvat), [Online]. Available: https://github.com/cvat-ai/cvat?tab=readme-ov-file#license.

УДК 004.8

UAeroNet: СПЕЦІАЛІЗОВАНИЙ НАБІР ДАНИХ ДЛЯ АВТОМАТИЗАЦІЇ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ

Юрій Кочура
http://orcid.org/0000-0002-4217-8152

Євгеній Трочун
https://orcid.org/0000-0002-2744-6681

Владислав Таран
http://orcid.org/0000-0003-2493-7239

Юрій Гордієнко
http://orcid.org/0000-0003-2682-4668

Олександр Роковий
http://orcid.org/0000-0001-6934-7502

Сергій Стіренко
http://orcid.org/0000-0001-5478-0450

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна

У цій статті розглядаються ключові аспекти та принципи розробки спеціалізованих наборів даних, які можуть використовуватися, зокрема для автоматизації безпілотних літальних апаратів. У дослідженні описано підходи, які ми використали для збору, аналізу та анотації даних, зокрема їхню значущість та практичне застосування у реальних умовах.

Підготовка спеціалізованого набору даних для автоматизації операцій безпілотних літальних апаратів (навігація, моніторинг довкілля) є складним завданням через часто низьку роздільну здатність зображень, складні погодні умови, великий діапазон масштабу об'єктів, фоновий шум та різнорідність ландшафтів місцевості. Наявні відкриті датасети зазвичай охоплюють лише обмежене різноманіття сценаріїв використання безпілотних літальних апаратів, що обмежує здатність моделей глибокого навчання адекватно працювати у нестандартних або непередбачуваних умовах.

Об'єктом дослідження є відеодані, отримані за допомогою безпілотних літальних апаратів для створення спеціалізованих наборів даних, які дозволяють моделям машинного навчання виконувати автономне розпізнавання об'єктів, навігацію, уникнення перешкод та взаємодію з навколишнім середовищем з мінімальним втручанням оператора. Предмет дослідження зосереджений на зборі, підготовці та анотації відеоданих, отриманих за допомогою безпілотних літальних апаратів. Мета дослідження – розробити та систематизувати робочий процес для створення спеціалізованих наборів даних, які можна використовувати для навчання надійних моделей, здатних автономно розпізнавати об'єкти в режимі реального часу на відео знятому за допомогою безпілотних літальних апаратів. Для досягнення цієї мети було розроблено робочий процес для збору та анотування відеоданих, зібрано сирі відеодані з датчиків безпілотних літальних апаратів та виконано ручну анотацію за допомогою Computer Vision Annotation Tool.

Розроблений нами спеціалізований набір даних для автоматизації безпілотних літальних апаратів в задачах відстеження об'єктів у практичних сценаріях отримав назву `UAeroNet`. `UAeroNet` складається з 456 анотованих треків та 131 525 розмічених екземплярів, що належать до 13 окремих класів.

Ключові слова: безпілотні літальні апарати, UAeroNet, виявлення об'єктів, автономна навігація, комп'ютерний зір.