

Novykov Danylo*National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv*

ORCID: 0009-0007-5373-6141

Onyshchenko Viktoriia*University of Warmia and Mazury in Olsztyn, Poland*

ORCID: 0000-0002-3126-2260

COMPUTER VISION-BASED APPROACH FOR MARKERLESS UAV LANDING ZONE IDENTIFICATION

Abstract: This article presents a computer vision-based approach for autonomous detection of safe landing zones for unmanned aerial vehicles (UAVs) using only images from an onboard camera. YOLOv5s was chosen as the base model, providing a good balance in speed and accuracy of detection with low computational complexity, allowing deployment in the resource-limited environments. The model was trained on a database of an urban environment containing four classes: “landing impossible”, “landing possible”, “person”, and “tree”. To increase the robustness, two data augmentation strategies were proposed that extend the input image processing pipeline at the model level. The first one utilizes the CLAHE, ToGray, or Equalize image augmentations, and another uses RandomBrightnessContrast, RandomShadow, or GaussNoise.

To check the adaptability of the proposed models to real-world variations, the evaluation session was conducted with different confidence thresholds: 25%, 50% and 75%. The results show that the modified models demonstrate a moderate improvement in key metrics. To further optimize results, an additional fine-tuning round was conducted using optimized hyperparameters and the weights from the initial stage. The results of evaluation highlight the efficiency of the proposed approaches.

Finally, based on the research results, the best model was selected for further use. Directions for future research are outlined, focusing on creating autonomous last-mile delivery systems using UAVs to increase the reliability and efficiency of delivery.

Keywords: object detection, image processing, unmanned aerial vehicles, UAV, autonomous landing, landing zone detection, autonomous delivery.

Новиков Данило Михайлович*Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ*

ORCID: 0009-0007-5373-6141

Онищенко Вікторія Валеріївна*Вармінсько-Мазурський університет в Ольштині, Польща*

ORCID: 0000-0002-3126-2260

ПІДХІД НА ОСНОВІ КОМП'ЮТЕРНОГО ЗОРУ ДЛЯ ВИЯВЛЕННЯ ЗОНИ ПОСАДКИ БПЛА БЕЗ ЗАСТОСУВАННЯ МАРКЕРІВ

Анотація: Метою дослідження є розробка моделі комп'ютерного зору для автономного виявлення безпечної зони посадки безпілотних літальних апаратів (БПЛА) за зображеннями зі вбудованої камери без використання спеціальних маркерів або додаткових систем, таких як GPS. В якості базової моделі обрано YOLOv5s як сучасний підхід до комп'ютерного зору, що поєднує гарну швидкість та точність детекції з невеликою обчислювальною складністю. Це дозволяє використовувати модель на різних кінцевих пристроях, навіть в умовах обмежених ресурсів. Базову модель було навчено на наборі тренувальних даних міського середовища, який містить чотири класи: “landing impossible”, “landing possible”, “person”, і “tree”.

Для підвищення стійкості до варіацій реального середовища запропоновано дві модифікації, які розширюють пайплайн обробки вхідних зображень на рівні моделі, збільшуючи варіативність даних під час тренування. Перша стратегія застосовує одну з трьох модифікацій із загальною ймовірністю 60%: CLAHE, ToGray або Equalize. Друга стратегія застосовує одну з наступних модифікацій зображення: RandomBrightnessContrast, RandomShadow або GaussNoise. Проведено раунд тренування моделей та отримано ключові показники, які стали основою для подальшого порівняння.

Наступним етапом перевірено адаптивність отриманих моделей до реальних умов, а саме проведено апробацію на тестовому наборі тренувальних даних, що містить трансформовані зображення. Для порівняння базової та модифікованих моделей проведено три етапи тестування з різними порогоми впевненості: 25%, 50% та 75%. Отримані результати наведено у відповідних таблицях. За результатами апробації модифіковані моделі демонструють стабільне поміrne покращення ключових метрик.

Для подальшої оптимізації та підвищення здатності моделей до узагальнення та покриття нестандартних ситуацій, проведено раунд донавчання з використанням оптимізованих гіперпараметрів на основі вагових коефіцієнтів з першого етапу. Результати проаналізовано та наведено висновки щодо ефективності запропонованих рішень.

У завершенні сформульовано висновки щодо доцільності застосування запропонованих підходів, обґрунтовано вибір найкращої моделі за результатами дослідження для подальшого використання. Окреслено напрямки майбутніх досліджень, спрямованих на створення систем автономної доставки останньої милі з використанням БПЛА для підвищення надійності й ефективності процесу доставки.

Ключові слова: виявлення об'єктів, обробка зображень, безпілотні літальні апарати, БПЛА, автономна посадка, виявлення зони посадки, автономна доставка.

1. Introduction

Unmanned aerial vehicles (UAVs) have evolved from expensive experimental proof-of-concept prototypes to modern devices widely used in everyday life. This is especially noticeable in the field of content creation, where drones offer new perspectives and opportunities for capturing materials. However, it is only a matter of time before this technology scales up and becomes a noticeable player in other areas, and logistics represents a field with especially high potential for UAV integration.

In the modern world, logistics is becoming increasingly complex due to the constant growth in the order volumes that require fast and reliable delivery. Integrating UAVs into the last-mile delivery allows the freeing up of human resources for more complex tasks that cannot be automated and require manual intervention, as well as improve the level of service for end users by reducing the time and cost associated with delivering small packages. An autonomous UAV delivery system can be considered as a combination of the following modules:

- Navigation module – responsible for planning the flight route and controls the movement of the UAV;
- Delivery endpoint determination module – helps identify the delivery location (evaluates whether it is the expected location, whether it is safe, and whether delivery is possible);
- Cargo control module – ensures the secure fixation and release of the cargo using the appropriate mechanism.

This work focuses on the delivery endpoint determination module – an important element that solves the task of selecting the endpoint in the delivery chain. Currently, there are no systems capable of determining a safe landing zone in real time within a typical urban environment for various scenarios (including private households and open public spaces such as parks and squares) without additional external dependencies (such as GPS or special markers).

2. Purpose of the study

The purpose of this study is to identify suitable landing zones for UAVs without relying on specialized markers. To achieve this, a computer vision-based model was developed to detect safe landing sites in open areas of public or private spaces that offer sufficient free space, using images from onboard cameras. Instead of using artificial markers, the model was trained to recognize natural surface features – such as smooth areas and the absence of obstacles. This approach improves the model's generalization and removes the need for artificial markers, making it more robust and applicable in real-world deployment.

3. Related work

A review of recent studies show that UAV-based autonomous delivery systems are advancing quickly, yet many persistent challenges remain. In earlier work [1], we introduced a computer vision model built on YOLOv5s, designed to identify safe landing zones for drones by analyzing images captured from onboard cameras. While this method proved effective in detecting landing zones marked with specialized visual markers under varying image distortions, its effectiveness is limited by the necessity of these markers being present. As a result, property owners are required to place markers in advance, and if no markers are available, like in a yard or a field that hasn't been prepared, the drone won't be able to find a safe place to land.

As another example, Ge et al. [2] propose a vision-based landing strategy for UAVs that enables docking on a custom multi-level platform under adverse wind conditions by detecting special markers – AprilTags. However, the solution is highly dependent on artificial markers and a specially designed landing surface, which limits its applicability in unprepared or natural environments. In addition, the article does not address important aspects of the model's generalization ability, such as its performance under varying lighting conditions and image distortions.

Commercial services have also tried to solve the problem of safe, flexible UAV delivery. Amazon Prime Drone Delivery [3] is a well-known example and has made real progress in automating parcel delivery. However, the user still has to pick a landing spot by pointing it out on a satellite photo of their property, so everything depends on how accurate the GPS is – and it means an extra step for the customer. Additionally, the UAV does not land, but drops the package from a height of 3.5 m, limiting the type of cargo, making the system less suitable for fragile or sensitive items, and reducing the versatility of the solution.

Flytrex [4], which currently operates in a few areas of Texas and North Carolina, offers another commercial solution. Customers can choose a preferred drop-off location (such as a backyard or workplace) using a mobile application. However, every new delivery point must first be approved in person by a company representative, who evaluates whether there is enough open space and checks for possible hazards such as overhead power lines or any other obstacles. In addition, the system is coupled to GPS coordinates and delivers packages using a zip-line mechanism from a height of 24–25 m, which complicates integration in dense urban environments and, similar to Amazon example, limits the types of cargo that can be safely delivered.

Taken together, well-known methods that depend on predefined landing markers or accurate GPS navigation encounter significant challenges – particularly in dynamic or GPS-denied environments. This highlights the potential of computer vision techniques to address the problem of safe landing zone detection by analyzing natural environmental “markers” such as surface flatness, absence of obstacles, and the availability of sufficient open space. By eliminating the dependency on artificial markers, these solutions offer the potential to boost the adaptability and effectiveness of future UAV delivery systems.

4. Results and discussion

In this study, YOLOv5s [5] was selected as the base model due to its well-established reputation as a state-of-the-art real-time object detection and image segmentation framework that offers high detection speed and accuracy, making it well-suited for resource-constrained environments. The training process utilized a dataset of urban environments sourced from the Roboflow platform [6]. The dataset does not contain any specialized landing markers. Instead, the focus of this research is on identifying safe landing zones within random urban environments using natural scene markers. The dataset includes four classes: “landing impossible”, “landing possible”, “person”, and “tree”. In total, there are 5773 labeled instances – 1479 labeled as “landing impossible”, 2090 as “landing possible”, 702 as “person”, and 1502 as “tree”. Utilizing this dataset, the model learns to recognize safe and dangerous landing zones and environmental obstacles.

During the first training round, standard hyperparameters were used to collect baseline metrics across all classes: Precision = 0.796, Recall = 0.7, mAP@0.5 = 0.726, and mAP@0.5:0.95 = 0.523. These results confirm the effectiveness of the baseline configuration when applied to urban images, demonstrating high initial detection capability under controlled conditions.

To improve the model's resilience to variations of the real-world environments, two alternative data augmentation pipelines were developed using the Albumentations library [7]. The first, named as Augmentation-A (used in the modified1 model), focuses on enhancing contrast and color representation to increase the model's tolerance to differences in image quality. During training, one of three augmentations is randomly applied to each image with a 60% probability: CLAHE (Contrast Limited Adaptive Histogram Equalization), which enhances local contrast; ToGray, which converts the image to grayscale; and Equalize, which adjusts image brightness and contrast via histogram equalization. Training with this augmentation pipeline produced the following results across all classes: Precision = 0.763, Recall = 0.711, mAP@0.5 = 0.721, and mAP@0.5:0.95 = 0.52.

The second strategy, named as Augmentation-B (used in the modified2 model), extends the augmentation pipeline with additional transformations aimed to simulate a wider range of real-world environmental conditions:

- RandomBrightnessContrast adjusts the image brightness and contrast within $\pm 20\%$ range with a 50% probability, improving the model's robustness to variable lighting;
- RandomShadow, applied with a probability of 30%, simulates shadows helping the model generalize better to environments with irregular illumination;

- GaussNoise adds Gaussian noise with a variance between 10 and 50 at a 30% probability, enhancing model's tolerance to sensor noise.

Using the Augmentation-B strategy, the model achieved the following results across all classes: Precision = 0.778, Recall = 0.709, mAP@0.5 = 0.728, and mAP@0.5:0.95 = 0.523. These results indicate a modest but consistent improvement, particularly in mAP score, indicating enhanced resilience to complex and variable visual conditions.

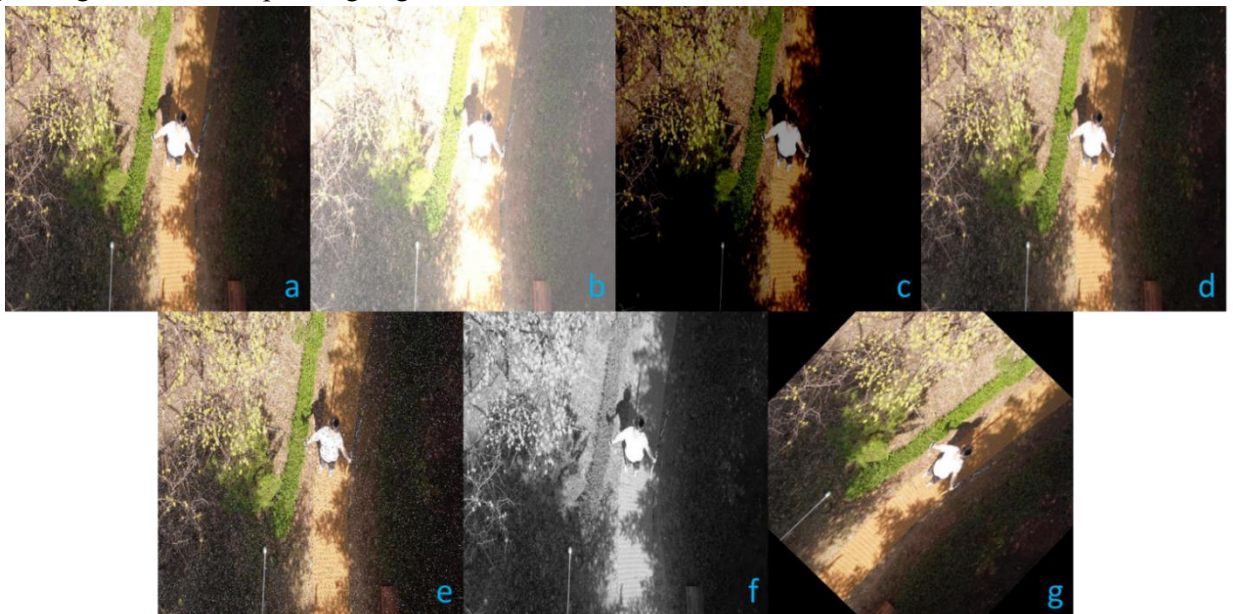
Table 1 presents the key metrics – both overall (across all classes) and for each individual class – obtained after training the baseline model and the models with extended augmentation pipelines (labeled as modified1 and modified2). A comparison of these two augmentation strategies demonstrates that integrating a wider variety of realistic transformations in the data pipeline can incrementally boost the model's overall detection quality and generalization ability, with the Augmentation-B strategy achieving the highest overall metrics as a result of the training session.

Table 1

Model performance after training					
Model	Class name	Precision	Recall	mAP@0,5	mAP@0,5:0,95
default	all	0.796	0.7	0.726	0.523
	landing impossible	0.799	0.726	0.768	0.57
	landing possible	0.698	0.52	0.541	0.35
	person	0.935	0.905	0.893	0.655
	tree	0.751	0.648	0.702	0.517
modified1	all	0.763	0.711	0.721	0.52
	landing impossible	0.747	0.737	0.754	0.562
	landing possible	0.65	0.528	0.535	0.352
	person	0.929	0.912	0.897	0.656
	tree	0.727	0.665	0.697	0.508
modified2	all	0.778	0.709	0.728	0.523
	landing impossible	0.737	0.726	0.753	0.563
	landing possible	0.682	0.521	0.54	0.357
	person	0.924	0.912	0.903	0.654
	tree	0.768	0.678	0.717	0.518

To evaluate the adaptability of both the baseline and modified models to real-world conditions, validation was conducted using a specially constructed test dataset [8], which contains a total of 1048 class instances. The test set includes: standard images (a), images with brightness adjustments of $\pm 50\%$ (b, c), Gaussian-blurred images (d), images with added noise (e), grayscale images (f), and images rotated by $\pm 45^\circ$ (g). Figure 1 shows a sample test image along with its corresponding augmented variants.

Fig. 1



Example of the test set with corresponding image augmentations.

Given that the primary goal of this work is to identify the optimal landing point by maximizing the probability of safe zone detection, the predicted region with the highest confidence score is selected as the final landing candidate for each test image. To compare the baseline and modified models, three phases of evaluation were performed – using confidence thresholds of 25%, 50%, and 75%.

Table 2

The results of the evaluation for the baseline (default) and modified (modified) models at a minimum confidence threshold of 25%

Model	Class name	TP	FP	FN	F_1	P	R	mAP@0,5	mAP@0,5:0,95
default	all	571	267	477	0.605	0.681	0.544	0.62	0.461
	landing impossible	149	86	131	0.579	0.634	0.532	0.623	0.478
	landing possible	148	126	229	0.454	0.54	0.392	0.44	0.327
	person	103	7	38	0.822	0.939	0.73	0.854	0.641
	tree	131	82	119	0.565	0.614	0.523	0.563	0.396
modified1	all	620	255	428	0.645	0.709	0.591	0.655	0.489
	landing impossible	174	89	106	0.641	0.661	0.621	0.675	0.508
	landing possible	162	121	215	0.491	0.573	0.43	0.483	0.36
	person	110	2	31	0.87	0.983	0.78	0.888	0.652
	tree	134	83	116	0.572	0.617	0.534	0.572	0.437
modified2	all	654	260	394	0.667	0.715	0.624	0.68	0.509
	landing impossible	185	96	95	0.659	0.658	0.661	0.71	0.56
	landing possible	181	105	196	0.546	0.632	0.48	0.516	0.378
	person	110	7	31	0.853	0.941	0.78	0.883	0.647
	tree	144	84	106	0.601	0.631	0.574	0.612	0.451

Table 3

The results of the evaluation for the baseline (default) and modified (modified) models at a minimum confidence threshold of 50%

Model	Class name	TP	FP	FN	F_1	P	R	mAP@0,5	mAP@0,5:0,95
default	all	450	142	598	0.549	0.761	0.43	0.603	0.466
	landing impossible	126	40	154	0.565	0.758	0.45	0.635	0.495
	landing possible	101	49	276	0.383	0.671	0.268	0.452	0.363
	person	99	2	42	0.818	0.981	0.702	0.846	0.643
	tree	75	44	175	0.407	0.632	0.3	0.477	0.365
modified1	all	494	135	554	0.589	0.786	0.472	0.632	0.494
	landing impossible	143	42	137	0.615	0.771	0.511	0.666	0.514
	landing possible	105	49	272	0.394	0.679	0.278	0.464	0.382
	person	107	0	34	0.863	1	0.759	0.879	0.656
	tree	85	38	165	0.455	0.692	0.339	0.518	0.424
modified2	all	503	135	545	0.596	0.788	0.48	0.642	0.501
	landing impossible	151	38	129	0.643	0.797	0.539	0.696	0.562
	landing possible	85	45	292	0.335	0.654	0.225	0.424	0.346
	person	110	3	31	0.866	0.974	0.78	0.887	0.651
	tree	94	35	156	0.494	0.727	0.374	0.563	0.446

For each phase, the following performance metrics were calculated:

- True Positives (TP): The number of correctly detected class instances;

- False Positives (FP): The sum of duplicate detections for a single class instance and incorrect predictions;
- False Negatives (FN): The number of class instances that were not detected.

Precision (1), Recall (2), and F_1 score (3), the harmonic mean of precision and recall, were calculated using standard formulas [9]:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F_1 = 2 \frac{Precision \times Recall}{Precision+Recall} \quad (3)$$

The evaluation results are summarized in Tables 2-4. For each model, results are provided across all classes as well as individually for each object class. Each block of results corresponds to a different minimum confidence threshold applied during detection: 25%, 50%, and 75%.

Table 4

The results of the evaluation for the baseline (default) and modified (modified) models at a minimum confidence threshold of 75%

Model	Class name	TP	FP	FN	F_1	P	R	mAP@0,5	mAP@0,5:0,95
default	all	338	70	710	0.465	0.829	0.323	0.579	0.467
	landing impossible	98	11	182	0.504	0.899	0.35	0.632	0.517
	landing possible	36	18	341	0.167	0.669	0.096	0.375	0.32
	person	94	0	47	0.8	1	0.666	0.833	0.638
	tree	45	15	205	0.29	0.75	0.18	0.478	0.395
modified1	all	373	80	675	0.497	0.823	0.356	0.593	0.486
	landing impossible	113	10	167	0.561	0.922	0.404	0.665	0.542
	landing possible	39	21	338	0.179	0.652	0.103	0.374	0.343
	person	104	0	37	0.849	1	0.737	0.869	0.651
	tree	45	18	205	0.287	0.718	0.18	0.465	0.41
modified2	all	372	58	676	0.503	0.865	0.355	0.612	0.504
	landing impossible	109	9	171	0.548	0.928	0.389	0.663	0.57
	landing possible	37	18	340	0.171	0.668	0.098	0.376	0.351
	person	100	1	41	0.827	0.99	0.709	0.853	0.63
	tree	56	8	194	0.355	0.873	0.223	0.554	0.466

Examination of results at different confidence thresholds highlights a consistent performance trend between the baseline and modified models:

- At a confidence threshold of 25%, both modified models show a noticeable improvement in the number of correctly detected instances (TP) compared to the baseline. For example, the overall TP for the baseline model is 571, while modified1 and modified2 achieve 620 and 654, respectively. This increase is reflected mostly by all classes, with the most noticeable difference in “landing impossible” and “landing possible” classes. Notably, modified1 outperforms the baseline by 8.6% in overall TP, and modified2 achieves a further 5.5% increase over modified1. Both augmented models also show higher key metrics (Precision, Recall, and mAP) scores compared to the baseline, indicating a better ability to generalize input data in the unpredictable conditions;
- At a confidence threshold of 50%, the advantages of the modified models become even more obvious as the threshold rises – modified1 achieves 494 overall TPs, while modified2 reaches 503, compared to 450 for the baseline. Although a slight decrease in recall is observed for certain classes, both modified models achieve higher overall key metrics. These results prove that the proposed augmentation strategies enhance the robustness of the models;
- At a confidence threshold of 75%, the number of detections naturally decreases due to the stricter confidence requirements. Nevertheless, both modified models continue to outperform the baseline model: modified1 records 373 overall TPs, while modified2 achieves 372, compared to only 338

for the baseline. In almost all cases, both modified models also maintain higher values across key metrics;

- This stable improvement in performance across different confidence thresholds is crucial for real-world applications where it is important to minimize the number of false positives.

In summary, both modified models demonstrate significant improvements over the baseline model at all tested confidence thresholds. While modified2 frequently achieves the best overall results, modified1 remains close and, in some cases, even outperforms modified2 in metrics such as Recall or F_1 score. These results highlight the importance of the proposed data augmentation strategies for improving the reliability and efficiency of UAV landing zone detection models, especially in varied and challenging environments.

To further enhance the models' generalization capabilities and performance in atypical scenarios, an additional round of fine-tuning [10] was conducted. This fine-tuning phase used optimized hyperparameters and was performed with the weights obtained from the initial training stage. The results of this process are presented in Table 5, where four configurations were evaluated:

- default*: the baseline model, fine-tuned from the initial weights without changing the hyperparameters (serving as a control version);
- ft-default: the baseline model, fine-tuned from the initial weights with newly selected, optimized hyperparameters;
- ft-modified1 and ft-modified2: the modified models with extended augmentation, fine-tuned from the initial weights using the optimized hyperparameters.

Table 5

Model performance after fine-tuning

Model	Class name	Precision	Recall	mAP@0.5	mAP@0.5:0.95
default*	all	0.788	0.698	0.719	0.519
	landing impossible	0.794	0.73	0.767	0.57
	landing possible	0.68	0.501	0.524	0.345
	person	0.951	0.905	0.902	0.661
	tree	0.726	0.655	0.683	0.501
ft-default	all	0.802	0.715	0.73	0.524
	landing impossible	0.811	0.722	0.754	0.565
	landing possible	0.71	0.541	0.573	0.382
	person	0.956	0.905	0.889	0.628
	tree	0.729	0.69	0.706	0.52
ft-modified1	all	0.79	0.726	0.731	0.515
	landing impossible	0.772	0.751	0.759	0.557
	landing possible	0.697	0.561	0.555	0.368
	person	0.953	0.891	0.884	0.607
	tree	0.74	0.701	0.727	0.529
ft-modified2	all	0.794	0.722	0.737	0.525
	landing impossible	0.763	0.74	0.747	0.564
	landing possible	0.685	0.535	0.567	0.384
	person	0.94	0.883	0.895	0.61
	tree	0.755	0.732	0.738	0.54

Compared to the baseline model, both ft-modified1 and ft-modified2 demonstrate noticeable improvements after fine-tuning. Specifically, ft-modified1 increased recall by 3.7%, while ft-modified2 improved recall by 3.1%. For mAP@0.5, ft-modified1 improved by 0.7%, and ft-modified2 by 1.5%. In terms of mAP@0.5:0.95, ft-modified2 achieved a 0.4% improvement, while ft-modified1 remained close to the baseline. Precision for both models remained nearly unchanged or showed a slight decrease, indicating that the primary benefits of proposed enhancements were in recall and average detection ability.

The obtained results indicate that the modified models, especially after fine-tuning, outperformed the baseline in most metrics, showing better generalization and reliability in challenging cases, confirming that the combination of advanced augmentation and hyperparameter optimization leads to more robust UAV landing zone detection across different scenarios. The fine-tuned models were also evaluated at various confidence thresholds using the test dataset, with results presented in Tables 6–8.

Table 6

The results of the evaluation for the fine-tuned models at a minimum confidence threshold of 25%

Model	Class name	TP	FP	FN	F_1	P	R	mAP@0,5	mAP@0,5:0,95
default*	all	549	220	499	0.604	0.714	0.524	0.619	0.447
	landing impossible	139	65	141	0.574	0.682	0.496	0.593	0.435
	landing possible	146	102	231	0.467	0.588	0.387	0.442	0.316
	person	105	10	36	0.819	0.911	0.745	0.854	0.626
	tree	117	57	133	0.553	0.673	0.469	0.587	0.412
ft-default	all	572	243	476	0.614	0.702	0.546	0.627	0.461
	landing impossible	140	81	140	0.559	0.632	0.5	0.599	0.451
	landing possible	142	106	235	0.455	0.574	0.377	0.46	0.354
	person	104	6	37	0.83	0.949	0.737	0.859	0.624
	tree	142	75	108	0.608	0.653	0.568	0.588	0.415
ft-modified1	all	613	247	435	0.643	0.713	0.585	0.654	0.471
	landing impossible	159	86	121	0.605	0.649	0.567	0.646	0.473
	landing possible	168	106	209	0.517	0.613	0.447	0.5	0.377
	person	105	7	36	0.83	0.939	0.745	0.859	0.596
	tree	146	78	104	0.615	0.651	0.583	0.611	0.437
ft-modified2	all	645	224	403	0.673	0.742	0.616	0.692	0.521
	landing impossible	171	68	109	0.659	0.715	0.611	0.706	0.548
	landing possible	174	96	203	0.538	0.644	0.462	0.531	0.401
	person	109	7	32	0.85	0.943	0.773	0.878	0.63
	tree	154	77	96	0.64	0.666	0.617	0.654	0.505

Table 7

The results of the evaluation for the fine-tuned models at a minimum confidence threshold of 50%

Model	Class name	TP	FP	FN	F_1	P	R	mAP@0,5	mAP@0,5:0,95
default*	all	465	125	583	0.568	0.788	0.444	0.613	0.457
	landing impossible	120	42	160	0.544	0.743	0.429	0.598	0.446
	landing possible	114	55	263	0.417	0.675	0.302	0.458	0.349
	person	100	3	41	0.821	0.974	0.709	0.847	0.627
	tree	84	26	166	0.465	0.76	0.335	0.549	0.407
ft-default	all	471	133	577	0.571	0.78	0.45	0.617	0.471
	landing impossible	123	45	157	0.549	0.732	0.439	0.608	0.469
	landing possible	115	56	262	0.419	0.673	0.305	0.472	0.38
	person	98	1	43	0.816	0.99	0.695	0.844	0.625
	tree	90	34	160	0.48	0.725	0.359	0.543	0.408
ft-modified1	all	541	154	507	0.621	0.778	0.517	0.648	0.48
	landing impossible	140	56	140	0.588	0.714	0.5	0.635	0.483
	landing possible	138	64	239	0.476	0.682	0.366	0.508	0.399
	person	103	3	38	0.834	0.973	0.73	0.858	0.597
	tree	118	41	132	0.576	0.743	0.47	0.591	0.443
ft-modified2	all	556	120	492	0.645	0.823	0.531	0.679	0.526
	landing impossible	150	34	130	0.647	0.817	0.536	0.7	0.556
	landing possible	129	47	248	0.466	0.732	0.342	0.515	0.412
	person	107	3	34	0.853	0.973	0.759	0.875	0.634
	tree	122	37	128	0.595	0.768	0.486	0.629	0.502

Table 8

The results of the evaluation for the fine-tuned models at a minimum confidence threshold of 75%

Model	Class name	TP	FP	FN	F ₁	P	R	mAP@0,5	mAP@0,5:0,95
default*	all	331	91	717	0.45	0.784	0.316	0.554	0.442
	landing impossible	90	24	190	0.457	0.787	0.321	0.57	0.449
	landing possible	33	31	344	0.15	0.519	0.087	0.298	0.262
	person	97	1	44	0.811	0.99	0.687	0.838	0.623
	tree	42	8	208	0.279	0.838	0.167	0.51	0.433
ft-default	all	327	73	721	0.451	0.818	0.312	0.567	0.461
	landing impossible	99	17	181	0.5	0.851	0.354	0.607	0.478
	landing possible	50	29	327	0.219	0.63	0.132	0.378	0.338
	person	81	0	60	0.729	1	0.573	0.787	0.609
	tree	47	13	203	0.303	0.789	0.188	0.496	0.418
ft-modified1	all	374	91	674	0.495	0.805	0.357	0.584	0.452
	landing impossible	114	20	166	0.551	0.851	0.407	0.638	0.496
	landing possible	60	27	317	0.258	0.688	0.159	0.413	0.35
	person	91	0	50	0.784	1	0.645	0.822	0.592
	tree	55	26	195	0.332	0.681	0.219	0.462	0.37
ft-modified2	all	369	83	679	0.492	0.817	0.352	0.591	0.484
	landing impossible	128	14	152	0.607	0.901	0.457	0.689	0.559
	landing possible	39	29	338	0.175	0.571	0.103	0.341	0.318
	person	88	0	53	0.768	1	0.623	0.812	0.615
	tree	56	14	194	0.349	0.797	0.223	0.521	0.445

Based on the evaluation of the fine-tuned models, the following conclusions can be drawn:

- Improvement in detection quality: both modified models demonstrate significant improvements in the number of TP and a decrease in FN detections across all test cases. Notably, ft-modified2 achieves up to 23.6% more TP and 17.7% fewer FN detections at the confidence threshold of 50%;
- Improved Recall and F₁ scores: enhancements in Recall and F₁ scores are especially notable, with an increase of up to 23.5% in Recall and 17.5% in F₁ score for the ft-modified2 model at the confidence threshold of 50%, indicating an overall boost in detection efficiency of the proposed model;
- Higher mAP values: the mean Average Precision (mAP) at both IoU thresholds shows a consistent improvement for both fine-tuned models. For instance, at the 25% confidence threshold, ft-modified2 improves mAP@0.5 by almost 9.7% and mAP@0.5:0.95 by 10.4% compared to baseline; at 50%, the improvements are 12.6% and 12.9%, and at 75%, 5.7% and 3.6%, respectively;
- Balanced Precision: although there are sometimes minor trade-offs in Precision at higher thresholds (75%), the overall increase in correct detections and mAP scores demonstrates improvements for real-world use;
- Best overall performer: among all tested configurations, the ft-modified2 model provides the best results across all metrics and thresholds, making it the recommended model for practical deployment.

Although the fine-tuned models overall outperform the default versions at lower and moderate confidence thresholds, their performance is slightly lower at the highest confidence threshold of 75%. This can be explained by the change in the distribution of confidence scores: fine-tuned models are less likely to assign very high confidence scores, which makes them more cautious in their predictions. As a result, some correct detections do not reach the target threshold level and are ignored, resulting in fewer TPs compared to the default models. In practice, such a cautious approach can help reduce FPs, even if it means missing a few correct detections at very high confidence levels.

5. Conclusion

In this work, we propose a vision-based approach for markerless UAV landing zone detection that uses only images from an onboard camera, without relying on specialized landing markers or GPS signals. The YOLOv5s model was adapted and trained to detect both safe and unsafe landing areas, as well as potential obstacles, using a realistic multiclass dataset. A series of experiments were conducted to evaluate the impact of advanced data augmentation strategies and hyperparameter optimization. Compared to the baseline, all proposed modifications – especially those involving data augmentation and hyperparameter tuning – provided a consistent improvement in detection quality, particularly for the critical class “landing possible”. Although these improvements are modest, they confirm the practical value of the proposed strategies for reliable UAV landing zone detection. Key findings include:

- Model-level data augmentation integrating a diverse set of realistic transformations during training not only increases the number of correctly detected safe landing zones, but also enhances the model’s overall reliability and robustness – key factors for deployment in dynamic and unpredictable environments;
- Hyperparameter optimization further enhances model performance compared to standard training;
- The integration of data augmentation and hyperparameter optimization provides the best balance between key metrics across all tested scenarios. The fine-tuned, augmented model (ft-modified2) achieved the highest combination of key metrics, including Precision, Recall, and F_1 score, outperforming the baseline and control models at most thresholds.

Based on the results, ft-modified2 is recommended as the most effective model for practical UAV landing zone detection. The proposed approach operates independently of GPS, making it suitable for environments where satellite navigation is limited or unavailable.

Future research should focus on expanding the current approach or integrating multiple models to enable the automatic recognition of specially equipped landing pads on multi-story buildings, ensuring coverage of all urban delivery scenarios. The application of the proposed model, in combination with navigation algorithms, opens up prospects for the development of fully autonomous, intelligent last-mile delivery systems using UAVs, which will significantly increase both the efficiency and reliability of aerial cargo delivery operations.

References

1. Novykov D., Onyshchenko V. Computer Vision-Based Approach for UAV Landing Zone Identification. *InfoCom Advanced Solutions 2025*, Kyiv, Ukraine, 21–22 May 2025. 2025. P. 123–125. URL: https://ist.kpi.ua/wp-content/uploads/2025/05/infocom-advanced-solutions-2025_compressed-3.pdf (access date: 28.06.2025).
2. Vision-Based UAV Landing with Guaranteed Reliability in Adverse Environment / Z. Ge et al. *Electronics*. 2023. Vol. 12 (4). URL: <https://doi.org/10.3390/electronics12040967>.
3. Gutierrez P. Amazon Prime Drone Delivery at Xponential Europe. *Inside Unmanned Systems*. 2025. 13 Mar. URL: <https://insideunmannedsystems.com/amazon-prime-drone-delivery-at-xponential-europe/> (access date: 05.07.2025).
4. Flytrex Home page. *Flytrex*. URL: <https://www.flytrex.com/> (access date: 05.07.2025).
5. Ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation / G. Jocher et al. *Zenodo*. 2022. URL: <https://doi.org/10.5281/zenodo.3908559>.
6. Landing-detection Computer Vision Project. *Roboflow*. 23.04.2025. URL: <https://app.roboflow.com/test-tqcr4/landing-detection> (access date: 14.06.2025).
7. Albumentations Documentation. *Albumentations*. URL: <https://albumentations.ai/docs/> (access date: 16.06.2025).
8. Landing-detection evaluation Computer Vision Project. *Roboflow*. 16.06.2025. URL: <https://app.roboflow.com/test-tqcr4/landing-detection-evaluation> (access date: 21.06.2025).
9. A Consolidated Overview of Evaluation and Performance Metrics for Machine Learning and Computer Vision / T. Schlosser et al. *ResearchGate*. 2024. URL: <https://dx.doi.org/10.13140/RG.2.2.14331.69928>.
10. Ultralytics YOLO Docs Insights on Model Evaluation and Fine-Tuning. *docs.ultralytics.com*. 29.06.2024. URL: <https://docs.ultralytics.com/guides/model-evaluation-insights/#how-does-fine-tuning-work> (access date: 18.06.2025).