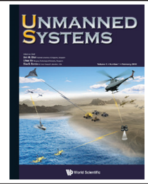
Unmanned Systems, Vol. 10, No. 1 (2022) 45–56

DOI: 10.1142/S2301385022500029

UNMANNED SYSTEMS
https://www.worldscientific.com/worldscinet/us



A Novel Framework for Multiple Ground Target Detection, Recognition and Inspection in Precision Agriculture Applications Using a UAV

Shahbaz Khan[*,†,¶], Muhammad Tufail[*,†], Muhammad Tahir Khan[*,†], Zubair Ahmad Khan[*], Javaid Iqbal[‡], Arsalan Wasim[§]

[*]*Department of Mechatronics Engineering, University of Engineering and Technology, Peshawar, Pakistan*

[†]*Advanced Robotics and Automation Laboratory, National Center of Robotics and Automation (NCRA), Pakistan*

[‡]*College of Electrical and Mechanical Engineering (CEME), National University of Sciences and Technology (NUST), Islamabad, Pakistan*

[§]*Department of Electrical Engineering, Hitec University, Taxila, Pakistan*

Unmanned Aerial Vehicles (UAVs) have been recently used for different civilian applications such as remote sensing, search, and rescue (SAR), precision agriculture (PA), etc. A UAVs ability to sense and find targets remotely and, based on that, hover close to the target for a particular action makes it an ideal platform for the aforementioned applications. There has been extensive work carried out in the field of visual-based detection, navigation, and control, but the problem of detecting different ground targets and performing certain actions is still an open research area. This study proposes a novel framework for multiple target detection, recognition, and navigation of the UAV to the desired target and closely inspect it. This proposed framework can be deployed for accurately spot spraying in PA applications or SAR. The target detection and recognition in the framework are achieved through a computationally efficient Convolutional Neural Network (CNN) trained model, whereas the close inspection of the target is achieved through a PID-based tracking algorithm which ensures the UAV hover around the target for few seconds. The developed framework performed the desired objective in five stages employing Lawson's control theory of sense, process, compare, decide and act. The target detection and recognition in the framework were validated with the field experiment, while the entire framework was validated through a variety of simulation flights conducted in Gazebo and PX4. The experiments' results showed the versatility of the developed system to many complex missions where the targets are added or removed.

Keywords: Unmanned aerial vehicle (UAV); unmanned aerial system (UAS); convolutional neural network (CNN); precision agriculture (PA); spot spraying, PID.

US

1. Introduction

Over the past few years, UAVs' research interests have increased remarkably, particularly regarding multirotor aerial robots, due to their flexible maneuvering ability to solve the complex missions [1]. It is expected that the sales of drones will surpass US$ 12 billion by 2021 [2]. Advancement in computer hardware and processing algorithms like deep learning has led to UAV deployment in various remote sensing applications such as SAR, ecology, wildlife and PA [3–5]. Remote sensing can assist farmers in the context of precision agriculture (PA) by assessing crop yield, health

Received 23 September 2020; Revised 29 March 2021; Accepted 30 March 2021; Published 29 April 2021. This paper was recommended for publication in its revised form by editorial board member, Jinwen Hu.
Email Address: ¶*shahbazkhan@uetpeshawar.edu.pk*

and disease [6, 7]. Different sensors that include: thermal cameras, multispectral cameras, RGB cameras, NIR cameras are utilized in UAVs for collecting data and monitoring crop health [8–10]. Image processing and other computer vision techniques are applied to the images obtained remotely for making decisions in several remote sensing applications. These decisions are either carried offline or online while the UAV is flying [5, 11]. The main focus in aerial work has been developing robust controllers instead of a reliable system that can detect and act on those targets. This ability can be beneficial in an application such as PA. For example, a UAV can accurately detect pest-infested crops and apply targeted pesticides instead of broadcast spraying, which has severe repercussions on the neighboring crops and the environment [12]. This ability of UAV is also required in SAR applications for delivering emergency medicines and also closely inspecting the target for victims etc., and making decisions accordingly.

This study aims to develop an autonomous vision-based framework for accurately detecting a target and acting, using UAV for PA applications, and demonstrating its robustness to errors. One of the key constraints associated with such missions performed autonomously through UAV is that multiple processes run simultaneously (e.g. Detection, tracking, interaction, etc.), implementing functions of each component. Usually, UAVs are deployed with computers having limited computational capabilities. So, it is essential to develop a framework that uses different computationally efficient techniques for performing the desired task. A modular software system was employed for implementing the proposed framework, and test cases were utilized for demonstration. This study's contributions are summarized as: (1) computationally efficient learning-based technique for accurately detecting and recognizing the target. (2) Onboard flexible framework for detecting and acting multiple ground targets which are robust to any errors. (3) The proposed framework having the ability to track and interact with the target can be used for different PA operations such as weed control, crop health monitoring and yield estimation. (4) The proposed framework evaluated in close inspection of multiple targets in a PA application in a simulated environment and real field experiments.

The rest of the paper is organized as related work is introduced in Sec. 2. Section 3 describes the study's framework. Experiments and results are detailed in Sec. 4. Section 5 presents the discussion. Conclusion and future research directions are drawn in Sec. 7.

2. Related Work

Developing a framework for detecting and acting on multiple targets is similar to the previous work carried in vision-based landing, target tracking, inspecting infrastructures, observe orient decide and act (OODA) approach.

Vision-based landing through Unmanned Aerial Vehicle System (UAS) is an extensively explored research area, where UAS uses vision for detecting and guidance towards a landing pad. Single targets such as landing pads are predominantly utilized for detecting the landing area, and an error of up to a few meters from the landing pad is tolerable. To improve detection and control capability of UAS "H" type, specialized patterns are commonly employed [5, 13]. Different techniques such as Image-Based Visual Servoing (IBVS) and relative distance-based approach are usually adopted, but their abilities are limited for multiple targets, especially when the targets' shape and size are unknown and similar [14]. Multiple target tracking was explored [15], employing IBVS and different target colors for separating them. However, when all of the targets or their subsets were visually similar IBVS schemes happen to be limited. Many researchers studied the ability to track moving targets [16]. For example, Rohan *et al.* [17] implemented onboard CNN-based object detecting and tracking system employing SSD architecture. However, multiple targets were not studied, and in some cases, the object detector missed the targets. In [18] the unmanned ground vehicle (UAV) was tracked with its relative location, but this technique has a limitation in multiple targets. Zhang *et al.* [19] demonstrated tracking of a user-selected object using UAS with a fixed gimbaled camera. To track the selected object, the authors proposed Kernelized Correlation Filter (KCF) based visual tracking.

In all the researches mentioned above, the study is only limited to a single target. Vision-based Navigation of UAS is also widely explored research area [15, 20, 21]. In [22], navigation inside an apple orchard was studied for UAS. Missions of following water channels [23] and roads have also been studied. UAS has also been utilized for inspection tasks such as power lines [24, 25], and railway track [26]. In these studies, normal lines with continuous image features were present. However, this study examines the problem of detecting and acting on multiple discrete targets where targets are added without prior knowledge to the system. Visual control techniques have also been widely studied for inspecting railway semaphore [27] and pole [28] employing UAS for flying around a pole-like structure. But, in these studies detection and approaching for action on target was not considered. Systems for finding a single target and acting on it utilizing vision have been investigated [3, 5, 28]. Systems having the desired capability to find and act on multiple targets have also been demonstrated [29, 30], but they were limited only to simple red color targets, and doubts about detecting the targets were not examined. An autonomous framework was established [1] for executing complex SAR missions. The system was able to

January 5, 2022 1:22:39pm WSPC/284-US 2250002 ISSN: 2301-3850 2nd Reading

detect multiple targets and interact with them accurately. However, the system was computationally expensive.

In addition to the aforementioned research, which mostly utilized the control system approach for controlling the UAS, another method uses the cognitive approach of OODA [31, 32]. For example, it was observed in [33] that on implementing and comparing the two controllers for UAS, OODA showed better results due to its property of being more intuitive and extensible. Their work was only confined to the simple mission of choosing alternative waypoints. An autonomous vision-based framework was designed [2, 14] for UAS employing the OODA approach for multiple ground target finding and action. However, their detecting method, which is very important in our study, was not accurate and was dependent on assumptions that false detections were not always persistent. In other related work, UAS has been extensively proposed in PA applications [34–37] Several authors employed different techniques such as Object-based Image Analysis (OBIA) [38], machine learning [36] for mapping agricultural fields. In these studies, image analysis was implemented off-board utilizing a powerful processing system. However, this study aims to develop a framework to perform the desired task in real-time using an onboard processing system. An onboard processing system [5, 39] was established for weed spraying. Though, only a single target was studied in the research.

In summary, extensive work has been carried out using target finding and acting on it, but accurate detection of the target and acting on it with autonomous decision-making operations is still an open research problem, and this study aims to explore it.

3. Framework Description

The framework proposed for autonomous navigation is the main contribution of this study. A modular software system was employed for developing and implementing the framework. The proposed framework comprises four abstraction layers: Sensors interface system, target detection system, main system and autopilot board system. Figure 1 illustrates the software system architecture.

All of the systems briefly explained in the subsequent sections were developed in Python. The autopilot driver system uses PX4 and Mavros [40], which is available as a ROS package [41] and provides communication between ROS and autopilot. The main system performs tracking navigation and control as a Lawsons Control Model [42] of Sense, Process, Compare, Decide and Act loop.

3.1. *Sensor interface system*

This system's main components comprise two main sensors: a camera sensor and an external sensor (Barometer).

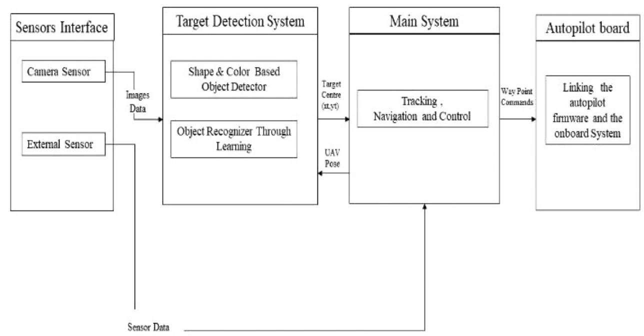


Fig. 1. Software system architecture.

The camera captures images and sends the image data into the target detection system. Additional estimated information of the state is provided to the main onboard system by an external sensor. In our proposed framework, a barometer is used to measure the altitude and achieve the desired height for hovering (inspecting) above the target.

3.2. *Target detection system*

The main aim of this system is to recognize and locate the target within the image plane. This system provides the target center utilized by the main system for tracking the target and performing the desired action of hovering above the target. This system comprises three main blocks to achieve the capability mentioned above, as depicted in Fig. 2.

3.2.1. *Target proposal using shape and color detection*

This module's main task in the system is to consider the computational constraint associated with aerial robots using an onboard processing system. This component uses low-cost-based algorithms to generate candidate proposals for the target recognizer module. This component enhances the effectiveness of the recognition stage, which is ultimately utilized for recognizing the targets. Moreover, the combination of shape and color components in the proposed algorithm can provide an intra-class classification of the recognized targets (e.g. crops, weeds, mud, plants, etc., as these targets differ based on shape and color). Additionally, this component is of great importance for

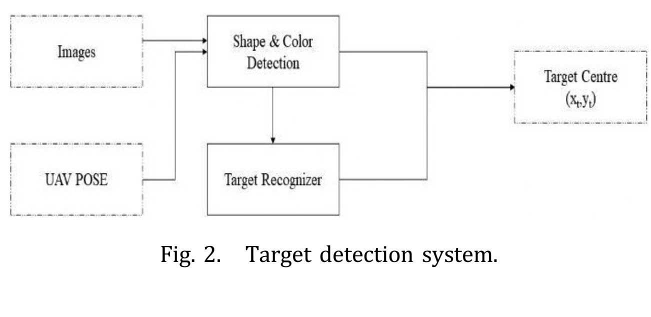


Fig. 2. Target detection system.

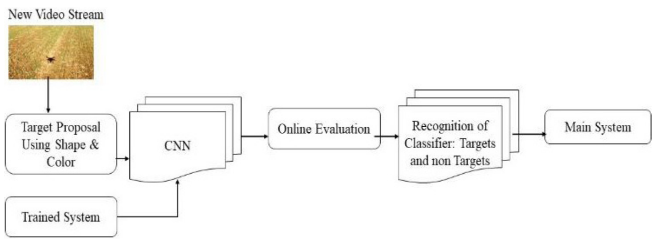


Fig. 3. Target recognizer.

computing the target center, as this component provides the contour of the target, which is utilized for calculating the target center and ultimately used for readjusting the Pose of the UAV.

3.2.2. *Target recognition*

This system component is responsible for accurately detecting the target and minimizing the vulnerability linked with the target detectors using predefined knowledge such as color, shape, features [14], etc. The study proposes a supervised learning classifier trained for target/background recognition, as illustrated in Fig. 3.

During this study, different supervised learning classifiers have been analyzed and explained in the experiment section.

After evaluating the experiment section, CNN architecture consists of five convolutional layers, four max-pooling layers, five dropout layers, two dense layers with one fully connected having a hidden layer of 512 units, and the last layer dense layer with sigmoid activation. The filter size is 3×3. $150 \times 150 \times 3$ is the input size of the image to the CNN model. RMS prop is selected as the cost optimization function.

3.2.3. *Target center estimation*

This module is responsible for computing the center of the target fed into the main system for the desired tracking of the target. This component finds the contours in the detected Region of Interest (ROI) in images obtained through target proposal using shape and color explained earlier. After finding the contours in the images, the image moments are utilized for detecting the center of the target. Weighted averages of the image pixels intensities define the image moments [43, 44].

Since we are studying binary objects in this study, the study targets are made of ones and backgrounds as zero. For binary image (p, q), moment and intensity $(i, j)M_{p,q}$ is given as

$$M_{p,q} = \sum_{i,j \in \mathrm{obj}} i^p j^q. \tag{1}$$

The total number of pixels in an object/target, i.e. object's area, is given by zero-order moment $M_{0,0}$. First-order moments $M_{1,0}$ and $M_{0,1}$ normalized by zero-order moments gives the desired targets center.

$$x_t = \frac{M_{1,0}}{M_{0,0}} \quad \text{and} \quad y_t = \frac{M_{0,1}}{M_{0,0}}. \tag{2}$$

3.3. ***Main system***

Lawson's model [42] is applied for implementing the main module of the proposed system. Figure 4 shows the flowchart system employing the concept of Lawson's model for onboard decision making. The system receives information about the environment through other modules based on which it decides and acts accordingly to achieve the desired task.

After the takeoff, the target detection algorithm recognizes and estimates the center of the target. Once the target is recognized, the UAV is updated to move towards the target. The UAV adjusts itself according to the target by aligning its x, y position so that the error between the target center and center of the current image becomes equal to zero [17] as defined as follows:

$$e_x(t) = x_t - x_i, \tag{3}$$

$$e_y(t) = y_t - y_i, \tag{4}$$

where x_t and y_t represent the target center and x_i and y_i represent the current image center.

After adjusting itself, the UAV descends according to the predefined height (1.3 m). When the target's height becomes equal to the predefined height, the UAV hovers above the target for 5 s. A control scheme based on PID is developed to readjust the UAV to descend and hover above the target, as illustrated in Fig. 5.

The proposed PID control scheme ensures that the parameters (roll, pitch, yaw and altitude) responsible for

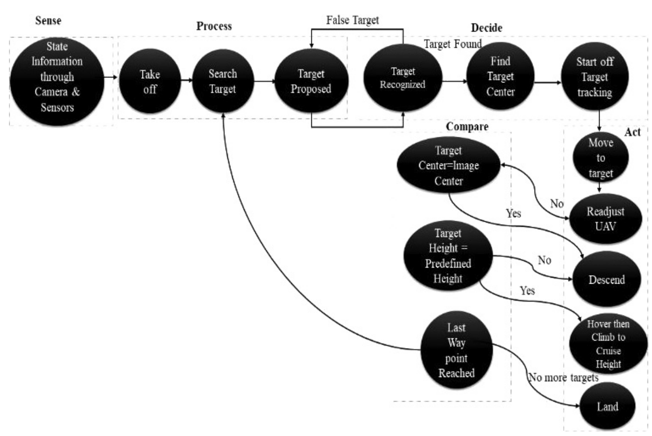


Fig. 4. Lawson's model flow chart for main system.

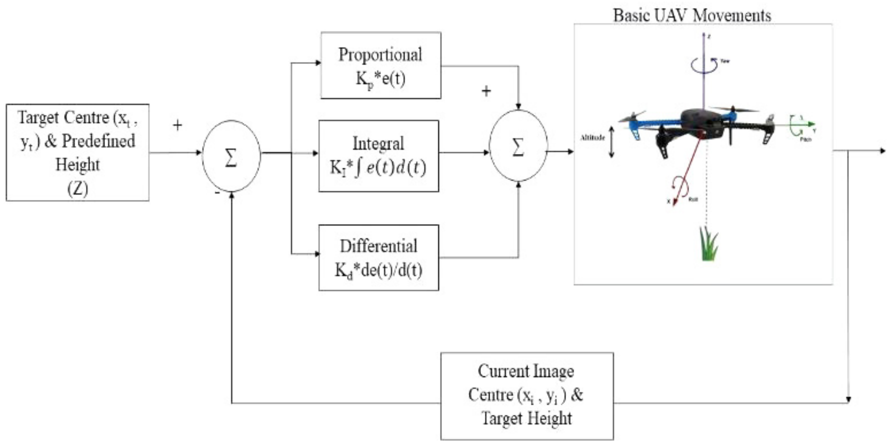


Fig. 5. Control scheme for main system.

the UAV movement achieve the desired task of readjusting, descending, and hovering above the target. Furthermore, the controller will help the UAV maintain its hovering position against external disturbances like wind. After this, the UAV either searches for a new target by climbing back to the cruising height or finishes the mission when all the targets are detected and acted accordingly.

3.4. *AutoPilot driver system*

This component utilizes PX4 autopilot for providing data such as autopilot state to the main system. Mavros Node is used for sending the waypoint commands to UAV for navigation. This component translates the commands received from the main system to its specific control commands.

4. Experimentation and Results

There are many safety and regulatory restrictions associated with the testing of UAVs in real-world outdoor environments. Simulations can be very beneficial to validate the proposed method and debugging the developed software code. In this regard, simulation experiments were conducted to validate the framework. The basic concept and typical mission of the framework are illustrated in Fig. 6.

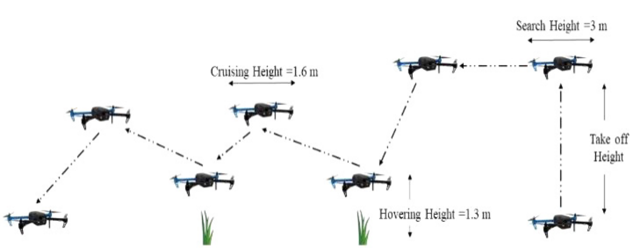


Fig. 6. Basic concept and typical mission for PA applications.

The mission can be divided into multiple steps, (1) Initially, the UAV takes off and starts searching for the target (search height = 3 m). (2) Once the target is detected and recognized, the UAV changes its path by moving towards the target and descends. (3) UAV hovers above the target ($h = 1.3$ m) for closely inspecting the target. (4) UAV climbs back to the cruising height ($h = 1.6$ m) and laterally moves to the next target. (5) UAV repeats the aforementioned steps until it reaches the point where there are no more targets. It lands and finishes the mission.

4.1. *Experiment scenario*

The scenario for the test case is illustrated in Fig. 7. The UAV flies with the height ($h = 3$ m) from waypoint A to waypoint B and search for targets. Once a UAV detects the target, it autonomously descends and hovers above every

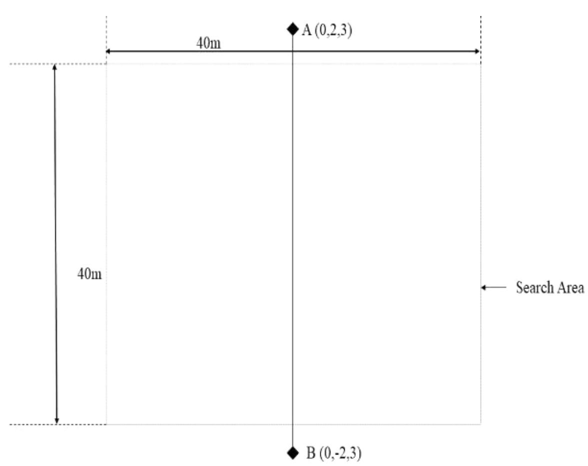


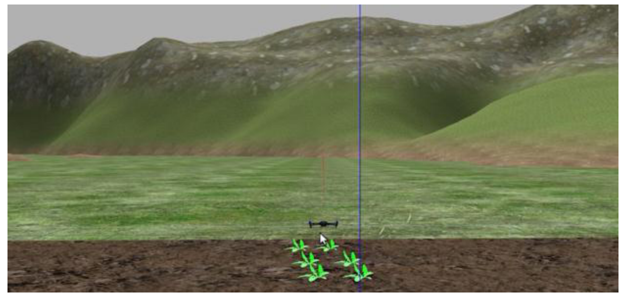
Fig. 7. Scenario for the test case. 40 m × 40 m represents the search area. A and B are the waypoints.

January 5, 2022 1:22:45pm WSPC/284-US 2250002 ISSN: 2301-3850 2nd Reading

target, as already explained in Fig. 6. Several green-colored targets in a row were placed between the waypoints.

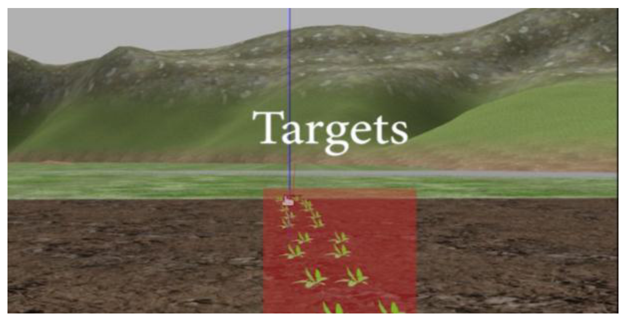
4.2. *Simulation experiment*

Different simulation approaches were explored in the study to yield realistic results, and the PX4 software in the loop (SITL) was selected for simulating the experiments because its simulation is nearly ideal. The environment was set up using a combination of the Gazebo robotic simulator and autopilot software stack. The environment was a single row of crops immersed in a muddy background. The targets were crops that have a standard shape as that of crops. The color of the crops was light green placed on mud. Target classification using a supervised learning classifier has been trained using Keras library. The platform selected for the simulations was an Intel i7-6600U processor, 8GB DDR4 memory, AMD Radeon R7 M360 2GB graphics card. Other plants were added along with the target in the environment, as shown in Fig. 8.



(a) Experiment

(b) UAV



(c) Target

Fig. 8. Experimental scenario where crops are used as targets.

4.2.1. *Training evaluation of supervised learning classifier for target recognition*

A custom data set containing targets (crops) and nontargets (noncrops) has been established. A total of 12,000 images for the crops class and 8000 images for the noncrops class were collected. The data set was divided into 70% for training, 15% validation and 15% test set. Three custom CNN models (see Table 1) varying the numbers of convolution layers, max pool and dropout layers and three pertained models VGG 11, VGG 16 and LeNet-5, were considered during the study for evaluation. A total of three evaluation tests have been defined to conduct a robust comparison, where random data sets were picked for training, validating and testing data sets. Keras Library having tensor flow at the backend was utilized for training the CNN configurations. Well-established methods such as VGG-11, VGG-16 and LeNet-5 were used for evaluation. The architecture of VGG-11 and VGG-16 consists of 8 and 13 convolutional layers, respectively, followed by three fully connected layers and a single SoftMax layer [45]. LeNet-5 comprises seven layers, consisting of two sets of pooling and convolutional layers, a single flattened convolutional layer, two fully connected layers, and at the end, a SoftMax layer [45].

After conducting three evaluation tests on the aforementioned classifiers, the average values of the obtained F1 results are listed in Tables 2 and 3.

It is evident from Table 2 that all the CNN configurations obtained a considerably high F1 score for the training set. Similarly, a significantly high score is achieved for the test sets for all of the configurations. The test time has been measured through simulations utilizing UAVs inbuilt system for calculating the processing time per image, for an average of 256 images, taking into consideration the feature extraction along with the time required for the desired classification.

Table 1. CNN configurations.

Architecture	Input dimension	No. of conv layers	No. of max pool layers	No. of dropout layers	Filter size	Feature map
CNN 1	$150 \times 150 \times 3$	3	2	2	3×3	32, 32, 64
CNN 2	$150 \times 150 \times 3$	4	3	3	3×3	32, 32, 64, 128
CNN 3	$150 \times 150 \times 3$	5	4	5	3×3	32, 32, 64, 128, 128

Table 2. Average training results of classifiers for the three evaluation tests.

	CNN 1		CNN 2		CNN 3		LeNet-5		VGG11		VGG16	
Target	F1 Score	Train time (s/epoch)	F1 score	Train time (s/epoch)	F1 score	Train time (s/epoch)	F1 score	Train time (s/epoch)	F1 score	Train time (s/epoch)	F1 score	Train time (s/epoch)
Crops	0.998	27	0.999	29	1.00	30	0.997	29	0.999	35	1.00	42
Noncrops	0.997		1.00		1.00		1.00		1.00		1.00	

Table 3. Average testing results of classifiers for the three evaluation tests.

	CNN 1		CNN 2		CNN 3		LeNet-5		VGG11		VGG16	
Target	F1 score	Test time (ms/epoch)	F1 score	Test time (ms/epoch)	F1 score	Test time (ms /epoch)	F1 score	Test time (ms/epoch)	F1 score	Test time (ms/epoch)	F1 score	Test time (ms/epoch)
Crops	0.923	3.79	0.941	3.82	0.978	3.90	0.937	3.22	0.985	4.65	0.988	5.84
Noncrops	0.921		0.938		0.972		0.930		0.984		0.986	

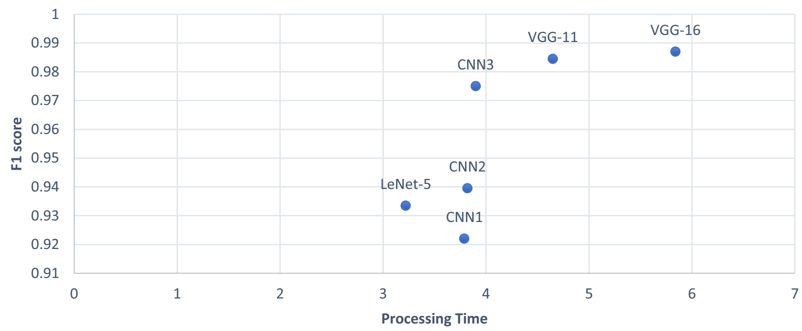


Fig. 9. Average F1 score and processing time for different configurations.

A fully autonomous aerial robot required for performing the desired mission of spraying PA or SAR requires multiple processes to be executed simultaneously. One of the main constraints associated with a UAV is its limited computational capability, making it essential to select an appropriate model to view the aforementioned constraint [1]. For this reason, in this study, like [1], the classifier selection is based on the F1 score (performance) and the test time (processing time), and Fig. 8 illustrates the plot of each configuration for better visualization.

After taking the results of Fig. 8 into considerations, the best possible classifier (minimum processing time and maximum F1 score) that has been selected for the study is CNN3, as it is the most suitable classifier for our purpose.

4.2.2. *Evaluation of the framework on the simulated environment*

A total of seven simulation experiments were conducted to validate the robustness of the developed system. The number of targets was increased in each test, and two vital parameters that are visiting the target and acting/hovering above the targets were analyzed. The results are summarized in Table 4. During all the tests, the developed framework was robust enough to visit and act on all the added targets even without knowing its addition, proving its re-planning ability. Furthermore, the targets were also removed in the waypoints, and the distance between the two consecutive targets was increased, the framework was able to

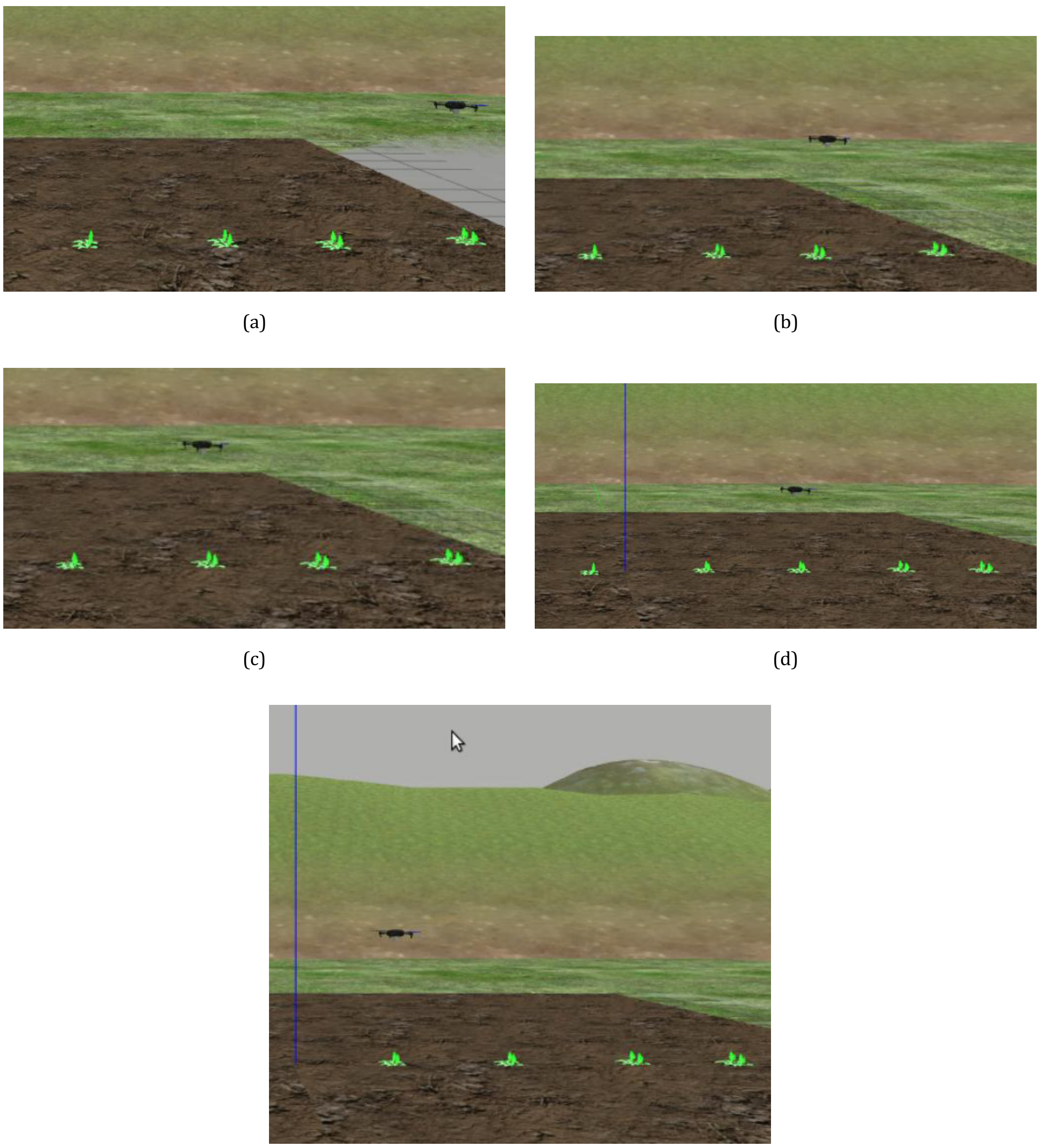
Table 4. Summary of simulation results.

Test number	Targets	Missed visiting targets	Missed hovering/ Acting targets
1	4	0	0
2	5	0	0
3	6	0	0
4	7	0	0
5	8	0	0
6	9	0	0
7	10	0	0

readjust accordingly and complete the mission. This showed its robustness which is of utmost importance in various PA applications (weed control, health monitoring, etc.). Figures 10(a)–10(e) shows visiting different targets and hovering above targets.

4.3. *Field experiment*

It was essential to perform a real-world experiment to validate the developed framework's target detection module. One-week flight tests were conducted at Turangzai (District Peshawar, KPK, Pakistan, Coordinates 34°12′57″ North, 71°44′50″ East). A quadcopter UAV was developed to perform the on-field experiments, as shown in Fig. 11, while the coriander field was considered for the study. The UAV was deployed with a Raspberry Pi4 onboard computer, camera, and intel neural computer stick 2 for onboard execution of the detection system. A height of 2 m was selected for acquiring images for training.



(a) (b) (c) (d) (e)

Fig. 10. Results of the UAV visited targets and interaction (a) Target 1 (b) Target 2 (c) Target 3 (d) Target 4 (e) Target 5.

January 5, 2022 1:22:52pm WSPC/284-US 2250002 ISSN: 2301-3850 2nd Reading

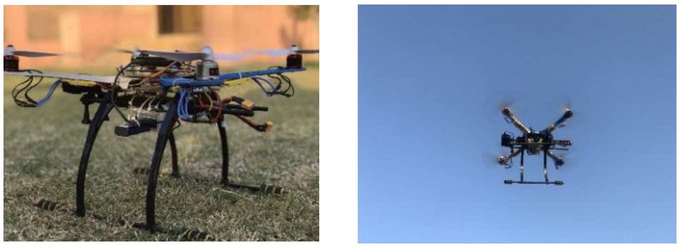


Fig. 11. Hardware in-flight tests.

Two classifier datasets (targets and nontargets) for coriander were collected for training the detection system. Images were obtained by converting the video in the pre-processing stage using a JPEG converter. A total of 2600 images for targets (coriander) and 1800 for nontargets were collected. The data was classified into training (70%), validation (15%), and testing (15%). The size of the input image was 448×448. Different lighting conditions were observed during the collection of the images during 1 week, and the average temperature was 20.2°C while the humidity was 58%.

4.3.1. *Evaluation of the target detection system*

Like the evaluation in a simulated environment, three evaluation tests have also been defined in the field experiments. Random data sets were chosen for training, validating, and testing. The developed classifier (CNN3) results and the three pre-trained models are summarized in Tables 5 and 6. The recognition of the targets (crops) is illustrated in Fig. 12. Recognition time has been measured through UAVs onboard computer for calculating process time per image (average of 256 images). The results were as follows (Table 6).

- F1 score of CNN3 was 0.945 with a recognition time of 4.13 ms.
- LeNet-5 achieved an F1 score of 0.905 and a processing time of 3.45 ms.
- F1 score for VGG-11 and VGG-16 was 0.95 and 0.975, while the recognition time was 5.31 and 7.2 ms, respectively

It is evident from the results that VGG-16 outperformed the developed classifier and other models based on the F1 score. However, the processing time, which is also an important parameter for UAV in VGG-16 is also higher. So, in this context, the developed classifier acts as an ideal choice for the developed target detection module of the framework.

5. Discussion

A UAVs ability to autonomously find multiple targets and perform particular actions (descend and hover) for inspecting the target closely, without human supervision, is an underdeveloped research area. This ability can be useful in several aforementioned applications such as SAR, PA applications, environmental monitoring, inspecting infrastructures, etc.

Finding and acting on multiple targets from a UAV is challenging because the system needs to keep track of the visited and acted upon targets. Usually, unique features such as shape, color, etc., of the targets are utilized for tracking. The study's main objective is to develop a robust

Table 5. Average Training results of classifiers for the three evaluation tests of field experiments.

	Developed classifier		LeNet-5		VGG11		VGG16	
Target	F1 score	Train time (s/epoch)	F1 score	Train time (s/epoch)	F1 score	Train time (s/epoch)	F1 score	Train time (s/epoch)
Crops	0.95	34	0.96	33	0.97	43	0.98	46
Noncrops	0.96		0.96		0.98		0.98	

Table 6. Average testing results of classifiers for the three evaluation tests of field experiments.

	Developed classifier		LeNet-5		VGG11		VGG16	
Target	F1 score	Test time (ms/epoch)	F1 score	Test time (ms/epoch)	F1 score	Test time (ms/epoch)	F1 score	Test time (ms/epoch)
Crops	0.94	4.13	0.90	3.45	0.95	5.31	0.97	7.2
Noncrops	0.95		0.91		0.95		0.98	

(a)

(b)

Fig. 12. Recognition of targets.

framework that can be applied in several applications where unique features cannot generally be ensured. For example, in the case of PA applications explained in the experiment section, it is likely to have multiple patches of crops and weeds having similar color, shape, and size. Detecting them and acting upon them is a challenging task.

The system developed in this study employs a flexible framework for performing a complex mission autonomously. The framework uses Lawson's control theory for achieving the desired task in a fully unsupervised manner. The developed system can automatically re-plan the tasks even when the targets are added or removed and hence can meet level 4 autonomy as per the NASA [46] evaluation method for autonomy. The system consists of an accurate target recognizer component based on CNN, keeping in view the computational constraints associated with a UAV. Furthermore, it consists of a PID-based control scheme for tracking and interacting with the target. The developed system has been extensively evaluated employing different simulated experiments, and the framework has been validated as evident from Table 4. One of our developed system's key capabilities is accomplishing complex missions and efficiently adapting to the changes that occur in the missions (e.g. recognizing additional targets).

The PA problem that has been addressed in this study by focusing on target detection, recognition and interaction is still an open field of research. The developed system has motivated future research that is summarized here:

(a) Validate the framework in real flight experiments. (b) Multiple target detection, recognition and interaction with targets in multiple rows. (c) Integrating obstacle avoidance capability into the framework to achieve full level 4 autonomy level. (d) Validate the framework's robustness by considering the errors and drifts in the onboard sensors and external disturbances such as wind.

6. Conclusion

This study developed a framework for autonomously finding and acting on multiple ground targets in a row using an UAS. The developed framework can be utilized in different UAV-based applications such as spot spraying in PA, search and rescue, etc. A framework was developed employing Lawson's control theory, different techniques and a modular software system for performing the desired task. Different supervised learning classifiers were evaluated for accurately recognizing the target and the final selected model of CNN consisted of five convolution layers. This exhibits a tradeoff between accuracy and the computational cost, keeping into account the hard-computational constraints of a UAV. The PID-based control scheme was adapted for tracking and interacting purpose.

Seven different test cases were developed for searching and inspecting multiple ground targets in a row. In each test, new targets were added and repeated several times to ensure the framework's robustness. The framework was able to adapt efficiently to the mission changes and was able to complete missions successfully. Further study on integrating the developed framework into real-time UAV is in progress.

Acknowledgments

This research was funded by the National Center of Robotics and Automation — Advanced Robotics and Automation Laboratory of UET Peshawar, Pakistan under an HEC grant.

References

[1] C. Sampedro, A. Rodriguez-Ramos, H. Bavle, A. Carrio, P. de la Puente and P. Campoy, A fully-autonomous aerial robot for search and rescue applications in indoor environments using learning-based techniques, *J. Intell. Robot. Syst. Theory Appl.* (2018) 1–27, doi: 10.1007/s10846-018-0898-1.

[2] Mohamed Natheer, Ajmal Hinas, A framework for vision-based ground target finding and action using multirotor UAVs, PhD diss., Queensland University of Technology, 2020.

[3] H. Choi, M. Geeves, B. Alsalam and F. Gonzalez, Open source computer-vision based guidance system for UAVs onboard decision making, *IEEE Aerospace Conf. Proc.*, Big Sky, Montana, USA, March 2016, pp. 1–5, doi: 10.1109/AERO.2016.7500600.

[4] S. Ward, J. Hensler, B. Alsalam and L. F. Gonzalez, Autonomous UAVs wildlife detection using thermal imaging, predictive navigation and computer vision, *IEEE Aerospace Conf. Proc.*, Big Sky, Montana, USA, 2016, pp. 1–8, doi: 10.1109/AERO.2016.7500671.

[5] B. Hazim, Y. Alsalam, D. Campbell, K. Morton and F. Gonzalez, Autonomous UAV with vision based on-board decision making for remote sensing and precision agriculture, *IEEE Aerospace Conf.*, Big Sky, MontanaT, USA, 2017, pp. 1–12.

[6] E. R. Hunt, M. Cavigelli, C. S. T. Daughtry, J. E. McMurtrey and C. L. Walthall, Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status, *Precis. Agric.* **6**(4) (2005) 359–378, doi: 10.1007/s11119-005-2324-5.

[7] H. Xiang and L. Tian, Method for automatic georeferencing aerial remote sensing (RS) images from an unmanned aerial vehicle (UAV) platform, *Biosyst. Eng.* **108**(2) (2011) 104–113, doi: 10.1016/j.biosystemseng.2010.11.003.

[8] C. C. D. Lelong, P. Burger, G. Jubelin, B. Roux, S. Labbé and F. Baret, Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots, *Sensors* **8**(5) (2008) 3557–3585, doi: 10.3390/s8053557.

[9] V. Gonzalez-Dugo *et al.*, Using high resolution UAV thermal imagery to assess the variability in the water status of five fruit tree species within a commercial orchard, *Precis. Agric.* **14**(6) (2013) 660–678, doi: 10.1007/s11119-013-9322-9.

[10] L. Felderhof and D. Gillieson, Near-infrared imagery from unmanned aerial systems and satellites can be used to specify fertilizer application rates in tree crops, *Can. J. Remote Sens.* **37**(4) (2012) 376–386, doi: 10.5589/m11-046.

[11] S. Von Bueren and I. Yule, Multispectral aerial imaging of pasture quality and biomass using unmanned aerial vehicles (UAV), Accurate and Efficient Use of Nutrients on Farms, Occasional Report 26 (2013), pp 1–5.

[12] M. D. Bah, A. Hafiane and R. Canals, CRowNet: Deep network for crop row detection in UAV images, *IEEE Access* **8** (2020) 5189–5200, doi: 10.1109/ACCESS.2019.2960873.

[13] S. Yang, S. A. Scherer and A. Zell, An onboard monocular vision system for autonomous takeoff, hovering and landing of a micro aerial vehicle, *J. Intell. Robot. Syst. Theory Appl.* **69** (2013) 499–515, doi: 10.1007/s10846-012-9749-7.

[14] A. Hinas, R. Ragel, J. Roberts and F. Gonzalez, A framework for vision-based multiple target finding and action using multirotor UAVs, *Sensors* **20**(1) (2020) 1320–1327, doi: 10.1109/ICUAS.2018.8453313.

[15] J. Zhang, Y. Wu, W. Liu and X. Chen, Novel approach to position and orientation estimation in vision-based UAV navigation, *IEEE Trans. Aerosp. Electron. Syst.* **46**(2) (2010) 687–700, doi: 10.1109/TAES.2010.5461649.

[16] M. T. Plehn, Control warfare: Inside the ooda loop, AiR UNIV MAXWELL AFB AL SCHOOL OF ADVANCED AIRPOWER STUDIES, 2000.

[17] A. Rohan, M. Rabah and S. H. Kim, Convolutional neural network-based real-time object detection and tracking for parrot AR drone 2, *IEEE Access* **7** (2019) 69575–69584, doi: 10.1109/ACCESS.2019.2919332.

[18] C. Greatwood *et al.*, Tracking control of a UAV with a parallel visual processor, *IEEE Int. Conf. Intelligent Robots and Systems 2017*, Vancouver, Canada, 2017, pp. 4248–4254, doi: 10.1109/IROS.2017.8206286.

[19] J. Ni, X. Zhang, P. Shi, J. Zhu and A. A. Salman, An improved kernelized correlation filter based visual tracking method, *Math. Probl. Eng.* **2018** (2018) 1–11, doi: 10.1155/2018/6931020.

[20] J. Zhang, W. Liu and Y. Wu, Novel technique for vision-based UAV navigation, *IEEE Trans. Aerosp. Electron. Syst.* **47**(4) (2011) 2731–2741, doi: 10.1109/TAES.2011.6034661.

[21] F. J. Perez-Grau, R. Ragel, F. Caballero, A. Viguria and A. Ollero, An architecture for robust UAV navigation in GPS-denied areas, *J. Field Robot.* **35**(1) (2017) 121–145, doi: 10.1002/rob.21757.

[22] N. Stefas, H. Bayram and V. Isler, Vision-based UAV navigation in orchards, *IFAC-PapersOnLine* **49**(16) (2016) 10–15, doi: 10.1016/j.ifacol.2016.10.003.

[23] A. S. Brandão, F. N. Martins and H. B. Soneguetti, A vision-based line following strategy for an autonomous UAV, *ICINCO 2015 — 12th Int. Conf. Informatics in Control, Automation and Robotics, Proc.*, Vol. 2, Colmar, France, 2015, pp. 314–319, doi: 10.5220/0005543903140319.

[24] O. Araar and N. Aouf, Visual servoing of a Quadrotor UAV for autonomous power lines inspection, *22nd Mediterranean Conf. Control and Automation, MED 2014*, Palermo, Italy, 2014, pp. 1418–1424, doi: 10.1109/MED.2014.6961575.

[25] S. J. Mills, J. J. Ford and L. Mejías, Vision based control for fixed wing UAVs inspecting locally linear infrastructure using Skid-to-Turn maneuvers, *J. Intell. Robot. Syst. Theory Appl.* **61**(1–4) (2011) 29–42, doi: 10.1007/s10846-010-9480-1.

[26] K. Máthé and L. Buşoniu, Vision and control for UAVs: A survey of general methods and of inexpensive platforms for infrastructure inspection, *Sensors* **15**(7) (2015) 14887–14916, doi: 10.3390/s150714887.

[27] K. Mathe, L. Busoniu, L. Barabas, C. I. Iuga, L. Miclea and J. Braband, Vision-based control of a quadrotor for an object inspection scenario, *2016 Int. Conf. Unmanned Aircraft Systems, ICUAS 2016*, Arlington, VA, USA, 2016, pp. 849–857, doi: 10.1109/ICUAS.2016.7502522.

[28] I. Sa, S. Hrabar and P. Corke, Inspection of pole-like structures using a visual-inertial aided VTOL platform with shared autonomy, *Sensors* **15**(9) (2015) 22003–22048, doi: 10.3390/s150922003.

[29] A. Hinas, J. M. Roberts and F. Gonzalez, Vision-based target finding and inspection of a ground target using a multirotor UAV system, *Sensors* **17**(12) (2017), doi: 10.3390/s17122929.

[30] A. Hinas, R. Ragel, J. Roberts and F. Gonzalez, A framework for vision-based multiple target finding and action using multirotor UAVs, *Int. Conf. Unmanned Aircraft Systems, ICUAS 2018*, Dallas (TX), USA, 2018, pp. 1320–1327, doi: 10.1109/ICUAS.2018.8453313.

[31] E. Theunissen, J. Tadema and A. A. H. E. Goossens, Exploring network enabled concepts for U(C)AV payload driven navigation, *AIAA/IEEE Digital Avionics Systems Conf. — Proc.*, Orlando, Florida, USA, 2009, pp. 1–15, doi: 10.1109/DASC.2009.5347463.

[32] S. C. Sari, Kuspriyanto and A. S. Prihatmanto, Decision system for robosoccer agent based on OODA Loop, *Int. Conf. System Engineering and Technology, ICSET 2012*, Bandung, Indonesia, 2012, pp. 1–7, doi: 10.1109/ICSEngT.2012.6339299.

[33] S. Karim and C. Heinze, Experiences with the design and implementation of an agent-based autonomous UAV controller, *Proc. Int. Conf. Autonomous Agents*, The Netherlands, 2005, pp. 59–66, doi: 10.1145/1082473.1082799.

[34] J. M. Peña, J. Torres-Sánchez, A. I. de Castro, M. Kelly and F. López-Granados, Weed mapping in early-season maize fields using object-based analysis of unmanned aerial vehicle (UAV) images, *PLoS One* **8**(10) (2013) 1–11, doi: 10.1371/journal.pone.0077151.

[35] M. Pérez-Ortiz, J. M. Peña, P. A. Gutiérrez, J. Torres-Sánchez, C. Hervás-Martínez and F. López-Granados, Selecting patterns and features for between- and within- crop-row weed mapping using UAV-imagery, *Expert Syst. Appl.* **47** (2016) 85–94, doi: 10.1016/j.eswa.2015.10.043.

[36] H. Huang, J. Deng, Y. Lan, A. Yang, X. Deng and L. Zhang, A fully convolutional network for weed mapping of unmanned aerial vehicle (UAV) imagery, *PLoS One* **13**(4) (2018), doi: 10.1371/journal.pone.0196302.

[37] S. Khan, M. Tufail, M. T. Khan, Z. A. Khan and S. Anwar, Deep learning based spraying area recognition system for unmanned aerial vehicle

based sprayers, *Turk. J. Elec. Eng. Comput. Sci.* **29** (2021) 241–256, doi: 10.3906/elk-2004-4.

[38] A. I. de Castro, J. Torres-Sánchez, J. M. Peña, F. M. Jiménez-Brenes, O. Csillik and F. López-Granados, An automatic random forest-OBIA algorithm for early weed mapping between and within crop rows using UAV imagery, *Remote Sens.* **10**(2) (2018) 1–21, doi: 10.3390/rs10020285.

[39] B. Alsalam, A small autonomous UAV for detection and action in precision agriculture. PhD diss., Queensland University of Technology, 2017.

[40] Ros.org, mavros — ROS Wiki. Available at: http://wiki.ros.org/mavros. Accessed 16 June 2020.

[41] Ros.org, Documentation — ROS Wiki. Available at: http://wiki.ros.org/. Accessed 16 June 2020.

[42] B. Brehmer, The dynamic OODA loop: Amalgamating Boyd's OODA loop and the cybernetic approach to command and control, *10th Int. Command and Control Research and Technology Symp. the Future of C2*, MacLean, Virginia, USA, 2005, December , pp. 365–368.

[43] L. Rocha, L. Velho and P. C. P. Carvalho, Image moments-based structuring and tracking of objects, *Brazilian Symp. Computer Graphic and Image Processing*, Fortaleza-CE, Brazil, 2002, January, pp. 99–105, doi: 10.1109/SIBGRA.2002.1167130.

[44] Raphaël, Tracking object orientation with image moments, (2016). Available at: http://raphael.candelier.fr/?blog=Image Moments. Accessed 19 June 2020.

[45] M. M. Rahman, M. S. Islam, R. Sassi and M. Aktaruzzaman, Convolutional neural networks performance comparison for handwritten Bengali numerals recognition, *SN Appl. Sci.* **1**(12) 1–11 (2019), doi: 10.1007/s42452-019-1682-y.

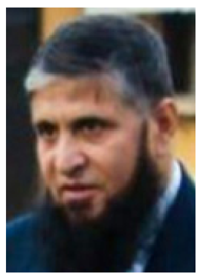
[46] L. A. Young, J. A. Yetter and M. D. Guynn, System analysis applied to autonomy: Application to high-altitude long-endurance remotely operated aircraft, in *Infotech@Aerospace*, Arlington, Virginia, 2005, p. 7103, doi: 10.2514/6.2005-7103.

Shahbaz Khan is a Ph.D. candidate in the Department of Mechatronics Engineering, at the University of Engineering and Technology (UET), Peshawar, Pakistan. He received his B.Sc. degrees in mechatronics engineering from UET in 2012, and his M.Sc. degree in Mechanical Engineering from Deakin University, Burwood, Australia in 2016. From 2017 to 2019, he was a Lecturer with the Department of Mechanical Engineering at Hitec University, Taxila, Pakistan. Since 2019, he has been a Lab Engineer at the Department of Mechatronics Engineering, UET. His research interests include agricultural robotics, control, and machine vision.

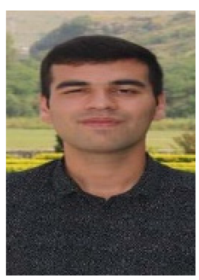
Muhammad Tufail received his B.Sc. degree in computer system engineering and M.Sc. degree in mechatronics engineering from the Asian Institute of Technology, Thailand, and his Ph.D. degree in mechanical engineering with specialization in manufacturing and mechatronics from the University of British Columbia (UBC), Canada. He then worked as a Postdoctoral Research Fellow with the Industrial Automation Laboratory, UBC, in 2016. Since 2017, he has been working as an Assistant Professor with the Department of Mechatronics Engineering, University of Engineering and Technology (UET), Peshawar, Pakistan. He is also the Co-Director of the Advanced Robotics and Automation Laboratory, UET, and a Co-PI with at the National Center of Robotics and Automation, Pakistan. His research interests include robot control, computer vision, and embedded systems. He is a member of IEEE.

Muhammad Tahir Khan received his master's degree in mechatronics from the University of New South Wales, Sydney, Australia, and his Ph.D. degree from the University of British Columbia (UBC), Vancouver, Canada, in 1999 and 2010, respectively. He was a Postdoctoral Fellow with the Industrial Automation Laboratory, UBC, for two years until January 2012. He is currently a Professor with the Department of Mechatronics Engineering at the University of Engineering and Technology, Peshawar, Pakistan. His research interests include robotics and intelligent control systems.

Zubair Ahmad Khan received his bachelor's degree in mechatronics engineering from the National University of Science and Technology, Peshawar, Pakistan, and his master's and Ph.D. degrees in mechatronics engineering from the University of Engineering and Technology (UET), Peshawar, Pakistan, in 2012 and 2019, respectively. He is currently a Lecturer with the Department of Mechatronics Engineering, UET. His research interests include manufacturing automation, robotics, and intelligent control systems.



Javaid Iqbal received his bachelor's and master's degrees in mechanical engineering from the College of Electrical and Mechanical Engineering (CE & ME), National University of Science and Technology (NUST), Islamabad, Pakistan, and Ph.D. degree from the University of New South Wales, Australia. He is currently performing his duties as the Dean of the Faculty of Engineering, CE & ME, NUST. His research interests include industrial simulation, machine learning, mechanical systems, and mobile robots, among others. He is a member of IEEE.



Arsalan Waseem Khan is currently pursuing his bachelor's in electrical engineering from HITEC University, Taxila, Pakistan. His interests are in the fields of deep learning, robot operating systems, and embedded systems.