WJCMS

Wasit Journal of Computer and Mathematics Science

Journal Homepage: https://wjcm.uowasit.edu.iq/index.php/WJCM

e-ISSN: 2788-5879 p-ISSN: 2788-5879

Computer Vision Techniques for Military Surveillance Drones

Hijaz Ahmad[1,*], Muhammad Farhan[2] and Umar Farooq[3]

[1]Section of Mathematics, International Telematic University Uninettuno, Corso Vittorio Emanuele II, 39,00186 Roma, Italy
[2]Institute of computing and information technology, Pakistan
[3]Department of Mathematics, Government College University Faisalabad, Faisalabad, 38000, Pakistan

*Corresponding Author: Hijaz Ahmad

DOI: https://doi.org/10.31185/wjcm.148
Received: April 2023; Accepted: June 2023; Available online: June 2023

ABSTRACT: Commercial drones, also known as unmanned aerial vehicles (UAVs), have grown recently, posing security problems and necessitating the development of efficient defenses. Radar, acoustics, and RF signal analysis are just a few of the technologies that have been investigated in order to overcome these issues. To identify autonomous drones, computer vision—in particular, deep learning techniques—has become a reliable and popular technique. This study aims to develop an autonomous drone identification and surveillance system using a combination of stationary wide-angle cameras and a lower-angle camera mounted on a rotating turret. We proposed a unique multi-frame DL identification model to optimize memory and processing time. In this method, the wide- angle still camera frames are placed on top of the frames taken by the turret's enlarged camera. With the help of this method, we can construct a productive pipeline that simultaneously conducts initial identification of small-sized aerial invaders on the primary picture plane and identification on the enlarged image plane. This method uses fewer resources and considerably lessens the computational load placed on detection algorithms. We also provide the entire system architecture, which comprises tracking algorithms, DL classification frameworks, and other critical elements. We build a complete drone identification and tracking solution by combining these components. To effectively classify and track drones in real-time, the system uses the power of deep learning, enabling quick responses and reducing possible security threats. Overall, this study provides a fresh and efficient method for autonomously identifying and tracking drones through the use of computer vision and deep learning algorithms. We offer a resource-effective approach to drone identification by integrating static and dynamic camera viewpoints with a multi- frame detection technique. This work supports continuing initiatives to strengthen security precautions against potential drone-related dangers.

Keywords: UAV or Drone, DL, Military Surveillance, and Computer Vision.

1. INTRODUCTION

Drones are increasingly being used in certain sectors, such as the military and surveillance, to make deliberate actions inside the battlefield. It is significantly more important to choose drones in order to provide real-time security. Real-time detection of drones in a variety of environments, including rain, sunshine, and night, is a major issue today. Deep learning is essential for identifying things in a variety of settings [1]. Faster R-CNN, R-CNN, and mask-RCNN are contemporary vision and DL algorithms that provide a technique to find an object [2].

The utilization of object identification and tracking systems has found wide-ranging applications, including military operations, healthcare, and security surveillance using autonomous robots [3]. Traditional object identification methods have primarily focused on edge detection, templates, and assumptions, often yielding lower accuracy and higher error rates [4]. Additionally, content-based image retrieval (CBIR) techniques have incorporated various feature extraction

*Corresponding author: Ahmad.hijaz@uninettuno.it

FIGURE 1. **Proposed Drone prototype**

methods, alongside filtering approaches, to ensure effective object identification [5]. Image classifiers have employed gradient-based histogram representations and local binary patterns within predefined object windows for efficient object recognition [6]. However, the implementation of these mechanisms within closed-circuit television (CCTV) systems using embedded platforms has posed challenges [7]. To address these limitations, deep learning techniques have emerged as a promising approach to improve accuracy [8].

DL models like Faster R-CNN, R-CNN, and Mask-RCNN are not well-suited for rapidly identifying tiny objects, including drones. In this study, a novel DL algorithm called UOSO (Unified Object Scale Optimization) is proposed to address this limitation [9]. To identify tiny drone objects, the suggested method employs a mix of confidence scores, multi-scale predictions, and an accurate classifier.

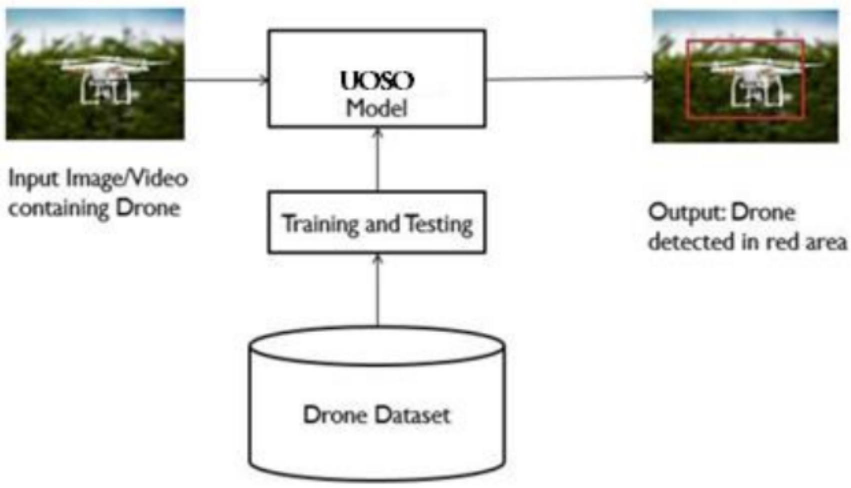

FIGURE 2. **Proposed method of this work**

For characteristic tiny drone objects, the urged deep UOSO uses the subsequent technique:

- the subsequent are the paper' contributions:
- A deep UOSO was plannedduring this paper to handle the tiny object identification downside quickly.
- The suggested deep UOSO model detects small drone objects with 99.99 % accuracy and token loss.

The remainder of the work is divided into five portions. The report summarizes the numerous literature reviews in Section 2 to sort out the issues. The deep UOSO model is proposed in Efficient Detection of tiny Drone Objects. The suggested deep UOSO model's simulation and performance to examine loss and accuracy are described in Section 4. Section 5 presents the findings and future directions.

2. LITERATURE SURVEY

According to recent developments, the utilization of industrial UAVs, basically referred as drones, has increased in various applications, including video and audio communication for security purposes [10]. In this context, a unique approach employing a multi-camera perspective is proposed for autonomous drone recognition and tracking [11]. The camera setup is thoroughly evaluated in terms of memory efficiency and processing time [12]. Specifically, small-sized aerial intruders are detected using compressed images and a deep learning-based detection algorithm [13, 14].

DL-based object identification methods have been suggested with varying levels of accuracy. For instance, Song et al. recommended the use of cost-effective aerial image with advanced drones, which initially posed challenges due to potential errors. They developed an embedded system framework for deep drones, where the drone's vision is comprehensively studied for autonomous tracking [15]. The embedded setup considers frame rate optimization, accuracy estimation, and power consumption analysis using the 1.6 Federal Protective Service for tracking [16]. YOLO (You Only Look Once) detectors, based on object detection and classification regression, have been introduced as a new approach [17]. Additionally, end-to-end performance detection employing neural networks with probability-based class analysis has gained popularity, with Faster R-CNN being extensively utilized for object identification [18].

Deep neural networks have been extensively used in computer vision to categorize and analyze binary and multimodal images. In addition to AlexNet, which is a popular network architecture with eight layers and sixty million connections, VGGNet has been introduced [8]. GoogleNet, utilizing convolutional layers with different scales and CNN assistance, has also been proposed [19]. This approach overcomes the gradient problem through cross-layer usage. ResNet, developed by He and his colleagues, improves image recognition accuracy by employing skip connections to bypass certain layers, while SqueezeNet achieves higher accuracy with fewer connections in CNN [20].

Kernelized correlation filters, which use the DFT (Discrete Fourier Transform) and a design with fast algorithms, have been used for picture categorization and detection. These filters achieve real-time tracking speeds, running at seventy frames per second on the NVIDIA TK1 kit [21]. Another study by Sabir Hossain et al. focuses on smart sensors and drones for target recognition and tracking in aerial images. They propose a deep learning-based framework using embedded modules like Jetson Texas or AGX Xavier with Intel Neural Computing Stick [22]. The accuracy of multi-object identification techniques using GPU-based embedded systems is assessed, considering the stipulated processing power [23]. DeepSORT accomplishes hypothesis tracking with the help of Kalman filtering by leveraging the association metric specified in multi-rotor drones. UAVs find applications in various civil and military domains, as highlighted by Roberto Opromolla, who explores visual cameras for tracking and detecting cooperative targets using frame sequences and deep learning [23].

Christos Kyrkou et al. provide a trade-off strategy for UAV identification in a specialized UAV environment utilizing a single-shot object detection model [24]. CNN offers a comprehensive improvement strategy with the use of remote-controlled aerial vehicles. Aerial images can be captured at speeds of 6-19 frames per second with 95% accuracy using low-power embedded processors. RetinaNet employs a backbone network for object detection, supporting both classification and regression tasks, while Faster R-CNN incorporates the Feature Pyramid Network (FPN) and merges input features from convolutional layers. These approaches utilize probability-based object presence with anchor-based mapping of input features in the pyramid levels [25].

The UAVs have been employed in power transmission devices, and DL algorithms have been utilized to control UAV transmission. Mask R-CNN has been employed for wireless communication, utilizing edge detection, hole filling, and Hough transforms.

3. PROPOSED METHODOLOGY

To achieve load balancing, various ML methods such as clustering and evolutionary algorithms have been presented. In the context of cognitive radio networks, a Markov model based on machine learning has been proposed for power optimization. YOLOv3 has been utilized to simulate different varieties of annotation errors in object identification, and the impact of these erroneous annotations has been analyzed during the training and testing phases. Furthermore, an embedded GPU-enabled flying robot equipped with a DL system has been designed to identify and track multiple objects in real time from aerial photographs. The proposed method employs kernelized correlation filters (KCF) of different sizes in different convolution layers. Overall, KCF proves to be exceptionally fast in video processing. The CNN layers split the picture into numerous areas and forecast the correct bounding boxes determined by the confidence scores for each separated region. The suggested UOSO algorithm is developed on a Dell EMC computer with Intel Xeon Gold 5118 12-core processors, NVIDIA Quadro GV100 GPUs, 256 GB 2666 MHz DDR4 ECC memory (six channels), and four 1TB NVMe Class 40 SSDs along with one 1TB SATA HDD.

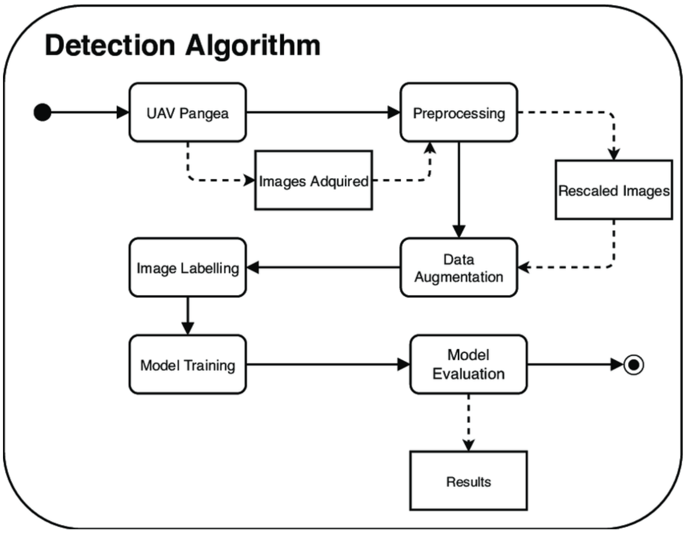

FIGURE 3. Flow chart of proposed method of this work

The proposed UOSO model represents a groundbreaking advancement in embedded deep learning for real-time detection of small objects. It adopts a single-instance approach to accurately predict the coordinates of bounding boxes for the whole picture. The model also determines the probabilities for each bounding box's object category. With an impressive frame rate of 45 frames per second, UOSO utilizes a deep CNN to achieve high prediction accuracy. Our suggested deep YOLO V3 prediction model for drones is shown in Figure 1. The training process involves 45,000 epochs using labeled input images. Each region in the model corresponds to a seven-by-seven grid, enabling the prediction of 5 bounding boxes. The suggested methodology effectively recognizes drone objects. The embedded model and the UOSO algorithm work together to forecast the categories and bounding boxes for the whole picture during a single run for a specified item location. Equation (1) defines the bounding box using four components: the center of the bounding box (by, bh), the dimensions (bw), the height (bh), and the object category (c).

$$y = (p_c, b_x, b_y, b_h, b_w, c) \tag{1}$$

For every bounding box, the CNN makes predictions for four coordinates: ty, tx, tw, and th. The following set of equations represents the process for predicting a bounding box. Cx and Cy are used to calculate the offset of an image's top-left corner, while pw and ph stand in for the preceding bounding box's width and height. During the training phase, the gradient ground truth value, denoted as t, is computed.

$$b_x = \sigma(t_x) + c_x \tag{2}$$

$$b_y = \sigma\left(t_y\right) + c_y \tag{3}$$

$$b_w = p_w e^{tw} \tag{4}$$

$$b_h = p_h e^{tk} \tag{5}$$

While a lot of of the grid and boxes don't embody a targeted item, the advised approach uses a non-max suppression technique termed freelance supplying regression to predict the pc. seven 7 grid cells are known at the same time victimization pc is to forecast a confidence score for a bounding box. the end result demonstrates that this is often} a extremely quick model. this method discards low-probability bounding boxes and predicts a bounding box with the very best probability. A solid secret score can be found within the anticipated bounding box. the boldness score is expressed in Equations (6).

$$P_r\,(o)\,x(IOU) \tag{6}$$

In the region, IOU is a crossroads for the union. It's calculated with the aid of using dividing the vicinity of the union of bounding containers with the aid of using the intersection vicinity. The floor reality field is stated to be 1 if IOU is amongzero and 1. The self belief rating for every bounding field that assures a field consists of an expected goal item is calculated the use of IOU. Background detection is likewise prevented with IOU. If there's no item with inside the grid cell, the self belief rating is zero. Otherwise, the self belief rating among the anticipated and floor reality bounding containers is identical to IOU. When the IOU is greater than zero.5, a higher prediction with excessive accuracy for item reputation is achieved. UOSO multiplies man or woman field self belief predictions with conditional magnificence chances said in Eq. 7 to get precise prediction.

$$P_r(c_i)IOU^{truth}_{prediction} \tag{7}$$

The confidence score of a bounding box is critical for generating a forecast during the testing stage. It is the result of a neural network.

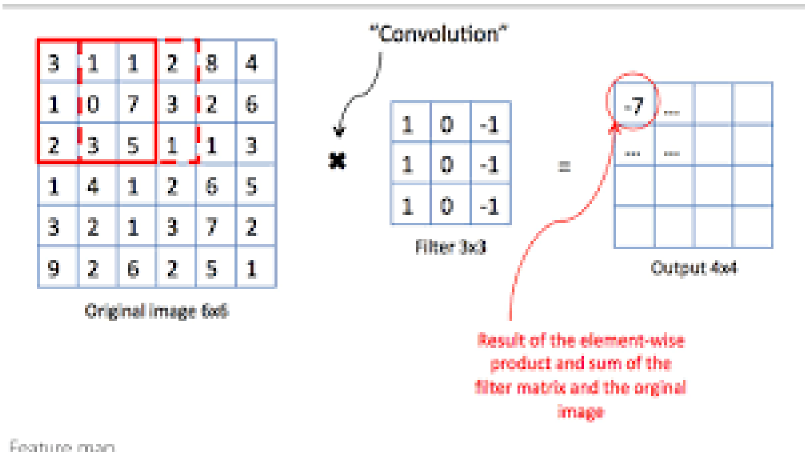

FIGURE 4. Future mapping diagram

The Deep CNN used in the research is made up of 106 convolution layers, comprising pooling, convolution, and fully connected layers with classifying abilities. By utilizing the input image and a sliding filter, we derived feature mappings, resulting in a two-dimensional matrix. Figure 2 illustrates an example feature map computed within the convolution layer.

4. RESULTS AND DISCUSSION

The study utilized 3000 drone pictures obtained from the Kaggle dataset, amounting to approximately two GB of Google data. A total of 45,000 epochs were performed on the drone dataset using the proposed UOSO approach, resulting in high sensitivity and accuracy. The example drone images that were utilized for practice and testing are shown in Figure 3. To facilitate the training process, a pre-trained UOSO model was employed. The implementation of UOSO was conducted on a GPU-based workstation.

A pre-trained UOSO technique with 6 convolution layers is employed to train the input images. Building the trained model takes quite eight hours throughout the training step. This UOSO model has been trained to just accept either image or video input. On the detection of drone photos or videos, the prompt model earned 99.99 p.c accuracy. within the testing stage, Figure four depicts the detection of drone video. Table 1 compares the accuracy of 3 models, together with UOSO, that is well matched for giant object detection in a very short quantity of time. as a result of it uses a hybrid network, the suggested UOSO designis suitable for tiny object identification.

The developed deep UOSO model, consisting of 106 convolution layers and multiple-sized feature maps, achieves an impressive accuracy of 99% during both the training and testing phases. By utilizing residual networks for object recognition, the YOLOv2 method is also employed for testing and training, achieving a success rate of 98.27%. To ensure

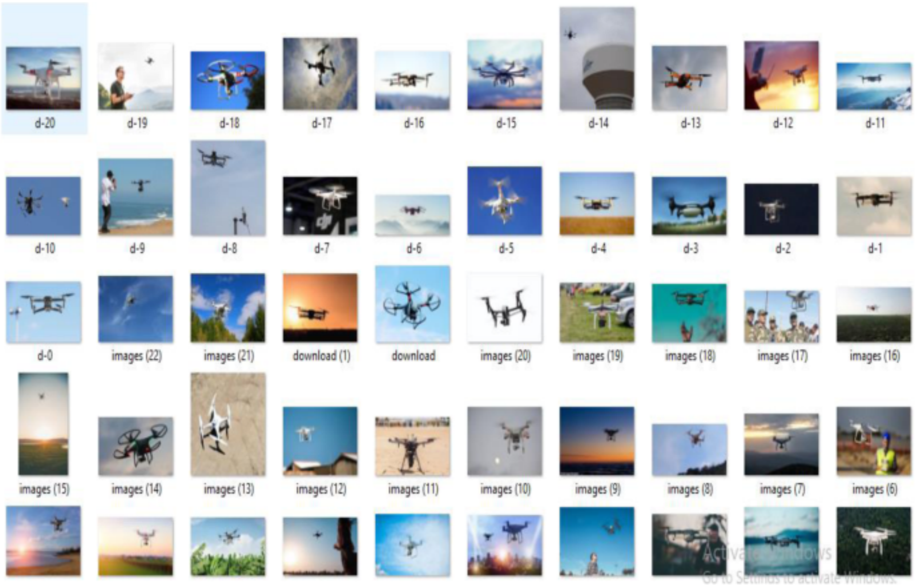

FIGURE 5. **Drone images**

Table 1. **Comparison of different models with existing and proposed method**

Model's	Accuracy (%)
YOLO [21]	85.37%
YOLOv2 [22]	90.60%
YOLOv3 [23]	95.10%
Proposed UOSO	99.96%

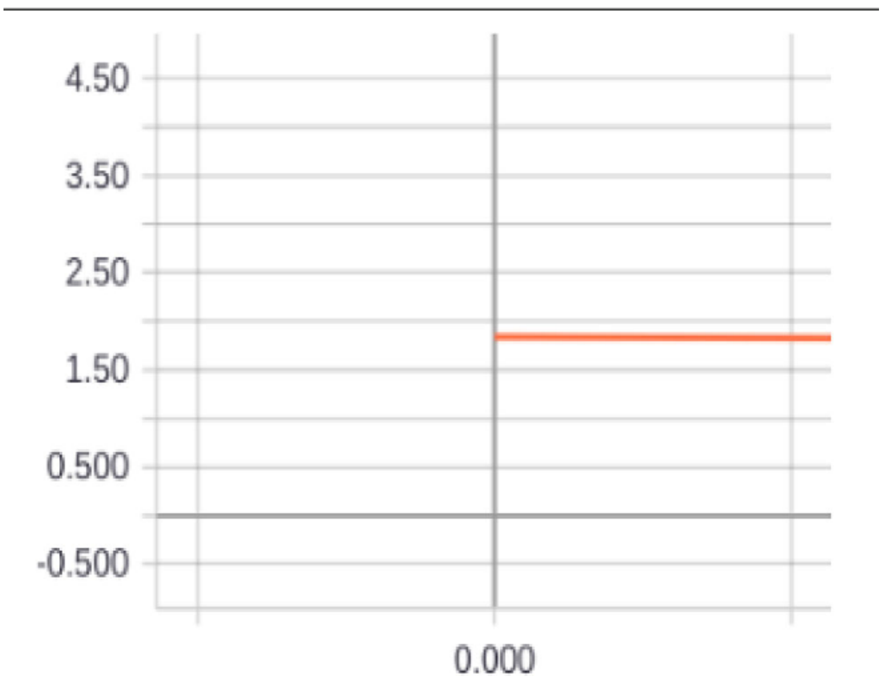

FIGURE 6. **Classification loss vs Epoch**

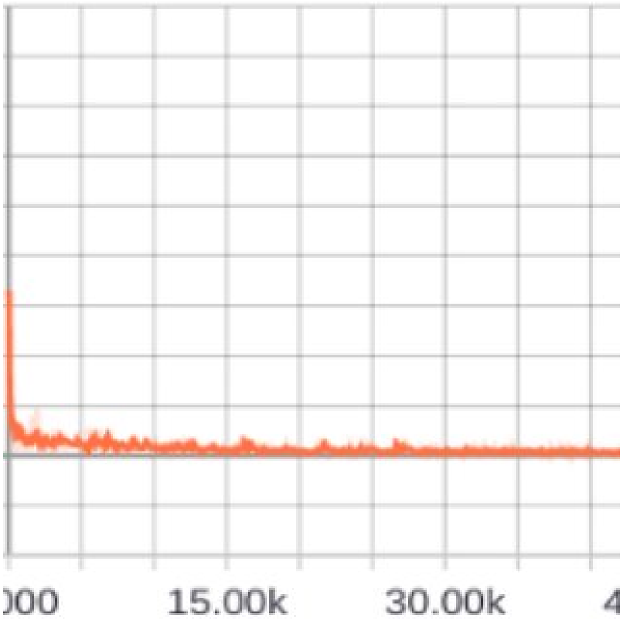

FIGURE 7. **Localization loss vs Epoch**

precise prediction of target objects, this method incorporates a confidence score relied on conditional probability. Figure 5 graphically illustrates the model's high accuracy after training. A backbone classifier is also included in the suggested technique to correctly identify the discovered items.

5. CONCLUSION

This paper presents a novel deep UOSO model specifically designed for the detection of small objects. The model is trained using drone images and leverages a pre-trained UOSO network. The simulation results show how well the suggested deep UOSO model performs in computer vision tasks. To effectively detect small drone objects, the model incorporates 106 convolution layers with diverse feature maps. By utilizing both Darknet and residual networks, UOSO is able to extract more informative features. The training process is performed for 45,000 epochs to ensure high accuracy. The recommended method employs a combined loss function to concurrently forecast the bounding box confidence score and the grid cell confidence score. To improve the recognition of tiny objects, the model also uses particular classifiers and binary cross-entropy loss optimization. As a result, the suggested deep UOSO model achieves an impressive accuracy of 99% by utilizing multi-scale predictions and backbone classifiers. The model also assigns a high confidence score to the detected objects based on conditional probability. Various loss functions are explored to optimize drone image prediction and accurately predict the precise boundary boxes of the objects. It should be mentioned that, in comparison to prior versions such as YOLO and YOLOv2, the proposed UOSO model is mainly geared for tiny items and may not be adequate for identifying bigger objects. Future work may involve expanding the algorithm to handle a larger volume of small drones in diverse visual conditions and remote locations.

FUNDING

None

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] F. L. Duarte and R. C. D. Lamare, “C-RAN-Type chd relaying with recursive maximum minimum distance,” *IEEE Communications Letters*, vol. 24, no. 11, pp. 2623–2627, 2020.

[2] L. Hu, Y. Tian, J. Yang, T. Taleb, L. Xiang, and Y. Hao, “Ready player one: UAV-Clustering-Based multi-task offloading for vehicular VR/AR gaming,” *IEEE Network*, vol. 33, no. 3, pp. 42–48, 2019.

[3] W. Yi, Y. Liu, E. Bodanese, A. Nallanathan, and G. K. Karagiannidis, “A unified spatial framework for UAV-a networks,” *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8801–8817, 2019.

[4] W. Feng, J. Wang, Y. Chen, X. Wang, N. Ge, and J. Lu, “UAVaided MIMO cit,” *IEEE Internet of 7ings Journal*, vol. 6, no. 2, pp. 1731–1740, 2019.

[5] H. He, S. Zhang, Y. Zeng, and R. Zhang, “Joint altitude and bo-enabled multiuser communications,” *IEEE Communications Letters*, vol. 22, no. 2, pp. 344–347, 2018.

[6] B. Li, S. Zhang, X. Zhang, B. Xi, and Y. Tian, “Application and research of scheduling mechanism for UAV cluster launching control system,” Xibei Gongye Daxue Xuebao,” *Journal of Northwestern Polytechnical University*, vol. 36, no. 2, pp. 353–358, 2018.

[7] W. Mei, Q. Wu, and R. Zhang, “Cellular-connected UAV: uplink association, power control and interference coordination,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5380–5393, 2019.

[8] Q. Zhu, K. Jiang, X. Chen, W. Zhong, and Y. Yang, “A novel 3D non-stationary UAV-MIMO channel model and its statistical properties,” *China Communications*, vol. 15, no. 12, pp. 147–158, 2018.

[9] X. Wang and M. C. Gursoy, “Coverage analysis for energyharvesting UAV-a cellular networks,” *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 12, pp. 2832–2850, 2019.

[10] M. Y. Arafat and S. Moh, “Localization and clustering based on swarm intelligence in UAV networks for emergency communications,” *IEEE Internet of 7ings Journal*, vol. 6, no. 5, pp. 8958–8976, 2019.

[11] D. Ebrahimi, S. Sharafeddine, P. H. Ho, and C. Assi, “UAVaided projection-based compressive data gathering in wireless sensor networks,” *IEEE Internet of 7ings Journal*, vol. 6, no. 2, pp. 1893–1905, 2019.

[12] S. Fu, Y. Tang, N. Zhang, L. Zhao, S. Wu, and X. Jian, “Joint unmanned aerial vehicle (UAV) deployment and power control for internet of things networks,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4367–4378, 2020.

[13] D. Liu, Y. Xu, and J. Wang, “Opportunistic UAV utilization in wireless networks: motivations, applications, and challenges,” *IEEE Communications Magazine*, vol. 58, no. 5, pp. 62–68, 2020.

[14] X. Liu, Y. Liu, and Y. Chen, “Reinforcement learning in multiple-UAV networks: deployment and movement design,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8036–8049, 2019.

[15] T. Yu, X. Wang, and A. Shami, “UAV-enabled spatial data sampling in large-scale IoT systems using da neural network,” *IEEE Internet of 7ings Journal*, vol. 6, no. 2, pp. 1856–1865, 2019.

[16] Y. S. Wang, Y. W. P. Hong, and W. T. Chen, “Trajectory learning, clustering, and user association for dynamically connectable UAV base stations,” *IEEE Transactions on Green Communications and Networking*, vol. 4, no. 4, pp. 1091–1105, 2020.

[17] M. Khabbaz, J. Antoun, and C. Assi, “Modeling and performance analysis of UAV-assisted vehicular networks,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8384–8396, 2019.

[18] D. Liu, J. Wang, and K. Xu, “Task-driven relay assignment in distributed UAV communication networks,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, 2019.

[19] W. Rahmaniar and A. E. Rakhmania, “Online digital image stabilization for an unmanned aerial vehicle (UAV),” *Journal of Robotics and Control (JRC)*, vol. 2, no. 4, pp. 234–239, 2021.

[20] A. Ibbotuwawa, G. Bocewicz, P. Nielsen, and B. Zbigniew, “Planning deliveries with UAV routing under weather forecast and energy consumption constraints,” *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 820–825, 2019.

[21] N. H. Motlagh, M. Bagaa, and T. Taleb, “UAV-based IoT platform: a crowd surveillance use case,” *IEEE Communications Magazine*, vol. 55, no. 2, pp. 128–134, 2017.

[22] X. Yu, X. Dong, and X. Yang, “Air-ground integrated deployment for UAV enabled mobile edge computing: a hierarchical game approach,” *IET Communications*, vol. 14, no. 15, pp. 2491–2499, 2020.

[23] X. Liu, Y. Liu, Y. Chen, and L. Hanzo, “Trajectory design and power control for multi-UAV assisted wireless networks: a machine learning approach,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7957–7969, 2019.