Investigating Domain Adaptation Feasibility for Drone Detection: A CPU-based YOLO Approach using RGB-Infrared Images

Warakorn Luangluewut
Dept. of Elect. Engineering
Kasetsart University,
Defence Technology Institute
Bangkok, Thailand
warakorn.l@dti.or.th

Phunsak Thiennviboon*
Dept. of Elect. Engineering
Kasetsart University
Bangkok, Thailand
phunsak.t@ku.th
*Corresponding Author

Kittakorn Viriyasatr
Defence Technology Institute
Bangkok, Thailand
kittakorn.v@dti.or.th

Piyarose Maleecharoen
Defence Technology Institute
Bangkok, Thailand
piyarose.m@dti.or.th

Teerasit Kasetkasem
Dept. of Elect. Engineering
Kasetsart University
Bangkok, Thailand
fengtsk@ku.ac.th

***Abstract*—Drone technology has become crucial in security and national defense, with anti-drone systems playing a vital role in government and military operations. A key component of these systems is drone detection using infrared camera imagery. While deep learning represents the state-of-the-art approach for object detection, it requires extensive datasets for practical implementation. Given the limited availability of infrared image datasets, leveraging larger RGB image datasets through domain adaptation could potentially enhance detection capabilities. This study investigates the feasibility of RGB-Infrared domain adaptation for drone detection, implementing CPU-based processing across various YOLO models (YOLOv5n/x, YOLOv10n/x, and YOLOv11n/x). We trained twelve models using either RGB or infrared datasets and evaluated their performance both with and without domain adaptation. Without domain adaptation, the models achieved excellent mean average precision (mAP50) values exceeding 95% at speeds of 0.18 – 10.65 frames per second (FPS). With domain adaptation, RGB-trained models detecting drones in infrared images achieved mAP50 values of 42.6 – 52.4% at 0.18 – 9.08 FPS, while infrared-trained models failed to detect drones in RGB images. Our findings demonstrate that (1) YOLO models excel at drone detection given sufficient data, (2) features learned from RGB images can be adapted for infrared image detection but not vice versa, and (3) domain adaptation with CPU-based processing is feasible for drone detection applications.**

Keywords—drone detection, domain adaptation, infrared image, RGB image, YOLO, CPU-based processing

I. Introduction

Drone technology has revolutionized numerous industries, particularly in aerial photography, surveillance, and military applications. The increasing prevalence of drones in modern warfare has created significant challenges for security and national defense. Anti-drone systems, designed to detect and intercept unmanned aerial vehicles in restricted or no-fly zones, have become crucial for government and military organizations. While these systems traditionally rely on human operators to evaluate potential threats, there is growing demand for automated detection capabilities.

Research in anti-drone technology has explored various detection methods [1-5], including radar-based systems [3] and object detection algorithms [7-11]. Deep learning models integrated with camera systems have shown particular promise [23], with significant advances in infrared-based drone detection [7-11]. However, effective infrared detection models require comprehensive datasets capturing multiple angles and scenarios. For instance, recent experiments implementing YOLO models in 5G systems utilized three infrared camera setups for effective drone detection [16-18]. The limited availability of infrared datasets has led researchers to investigate the potential of incorporating RGB image data [12-15].

This study investigates the feasibility of domain adaptation [6] for drone detection across imaging modalities, specifically examining whether models trained on RGB images can effectively detect drones in infrared images, and vice versa. We evaluate both detection accuracy and processing speed to assess practical deployment potential with real-world cameras. Our testing implementation focuses on CPU-based processing, following Bhattacharya's demonstration of its suitability for portable, cost-effective hardware solutions [7]. The investigation encompasses six YOLO architectures in both Nano (YOLOv5n, YOLOv10n, YOLOv11n) and Extra Large (YOLOv5x, YOLOv10x, YOLOv11x) configurations, to evaluate detection performance and processing speed trade-offs across established and recent model variants [19].

The remainder of this paper is organized as follows: Section II provides relevant background material, Section III details our methodology, Section IV presents results and discussion, and Section V concludes the study.

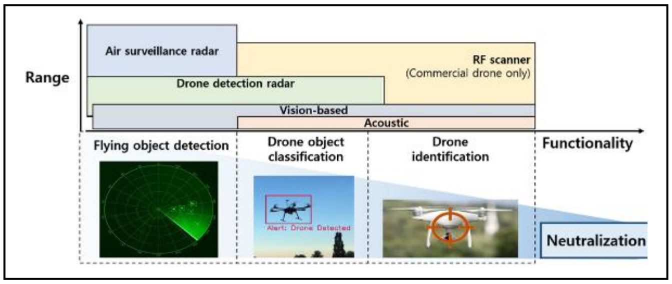

Fig. 1. Categorization of anti-drone technologies including their detection and classification methods with respect to functionalities and ranges [1].

II. Background

A. Anti-Drone Technology

Anti-drone systems, also known as counter-UAS (Unmanned Aircraft Systems), comprise technologies and strategies designed to detect, track, identify, and neutralize unauthorized or potentially threatening drones. The proliferation of drone technology and its diverse applications has heightened the importance of these systems. Anti-drone systems integrate multiple detection technologies, including radar, radio frequency (RF) monitoring, optical and acoustic sensors, and deep learning algorithms.

These systems operate through three primary components:

- Detection: Detection of drone presence within a designated area, utilizing technologies such as radar systems, radio frequency (RF) monitoring, and camera-based imaging systems with advanced image processing algorithms.
- Identification: Analysis of drone characteristics including flight patterns, physical attributes, and signal signatures to assess potential threats. This phase increasingly employs machine learning technologies, particularly Convolutional Neural Networks (CNNs), to achieve higher accuracies.
- Neutralization: Implementation of countermeasures against identified threats, typically through signal jamming or other neutralization methods.

Park et al. [1] provide a comprehensive survey on anti-drone systems. Figure 1 categorizes these technologies based on detection methods, functional capabilities, and operational ranges [1]. Despite significant technological advances, anti-drone systems continue to face challenges, particularly in distinguishing drones from other aerial objects like birds across diverse environmental conditions. These persistent challenges necessitate ongoing research to enhance system effectiveness in real-world deployments.

B. Drone Detection through Infrared Images

Infrared imaging captures electromagnetic radiation with wavelengths between visible light and microwaves, typically ranging from 750 nanometers to 1 millimeter [20]. This technology detects thermal radiation emitted by objects above absolute zero (-273.15°C), converting it into visible representations. Commercial infrared cameras, such as Sony Alpha 6000 [16], FLIR [17], and VarioCAM [18], offer frame rates between 9 and 60 FPS for thermal imaging applications.

Infrared imaging can be used for thermal analysis and low-light operation. Its ability to detect temperature variations enables to use infrared images as thermal mapping across surfaces and objects, while its independence from visible light allows effective operation in dark environments. These capabilities have proven particularly valuable for drone detection. where recent research [7-11] has been conducted using deep learning-based object detection models [23-24] for drone localization and classification in infrared images, as illustrated in Fig. 2.

C. YOLO (You Only Look Once) and Model Scaling

YOLO (You Only Look Once) [21, 22] is a popular family of deep learning models suitable for real-time object detection while maintaining high accuracy. The YOLO architecture reframes object detection as a regression problem, enabling simultaneous classification and localization through a single forward pass [22]. These models generate tensor outputs containing predicted bounding box information (locations, sizes, confidences) and conditional class probabilities of detected objects. Successive YOLO iterations have demonstrated improved performance on complex scenes while maintaining flexibility for various detection tasks. The YOLO framework has undergone continuous development, achieving better performance while allowing customization for specialized detection requirements. Model scaling [19] enhances detection capability through systematic adjustment of architectural dimensions. Typically, model scaling includes adjustments in depth and width where

- Depth scaling: Increases the network's layers, enabling the capture of more complex features through additional convolutional layers [26] in the neural network.
- Width scaling: Expands the number of units per layer, typically through additional convolutional filters, allowing broader feature detection at each layer.

D. Domain Adaptation

While classical machine learning relies on the assumption of identical distributions between training and testing datasets, practical applications frequently encounter different data distributions. These differences arise from various factors, including limited data availability demanding diverse sources, or temporal evolution of data characteristics. Domain adaptation [6] provides a framework for addressing these distribution mismatches, enabling models to effectively generalize across different domains. This study explores domain adaptation between RGB and infrared imaging modalities, leveraging their shared feature characteristics for object detection tasks.

Fig. 2. An example of a drone detected in an infrared image using a YOLO model.

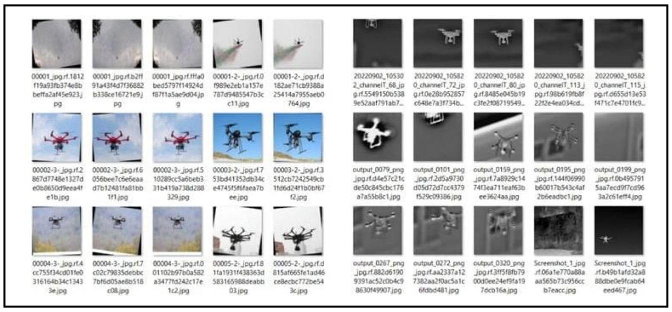

Fig. 3. Representative images from RGB (left four columns) and infrared (right four columns) drone detection datasets.

III. Research Methodology

This study investigates the feasibility of domain adaptation for drone detection in infrared images using RGB image datasets. We evaluated twelve YOLO models, conducting model training and model testing both with and without domain adaptation. Detection performance was systematically analyzed to assess the viability of domain adaptation in drone detection applications. The following sections detail our methodology, including dataset preparation, data partitioning, model selection, and performance metrics.

A. *Dataset Preparation*

This study employed two drone detection datasets: an infrared dataset comprising 7,060 labeled images [29] and an RGB dataset containing 7,792 labeled images [30]. Both datasets include bounding box annotations for drone localization. Representative samples from both datasets are shown in Fig. 3.

B. *Data Partitioning*

Each dataset was partitioned into three subsets for training, validation, and testing. Image augmentation techniques [31] were applied separately to each subset to enhance dataset size and diversity. Table I summarizes the distribution of RGB and infrared images, along with the number of drone instances across all augmented datasets.

C. *Model Selection*

We selected YOLO models for this study based on their demonstrated high accuracy and computational efficiency [22], crucial factors for real-time drone detection. The investigation encompassed six YOLO variants across three architectures: YOLOv5, YOLOv10, and YOLOv11. YOLOv10 and YOLOv11 were chosen for their recent technological advancements, while YOLOv5 was included due to its widespread adoption in drone detection research. For each architecture, we implemented both Nano (YOLOv5n, YOLOv10n, YOLOv11n) and Extra Large (YOLOv5x, YOLOv10x, YOLOv11x) scaling configurations, balancing the trade-off between processing speed and detection accuracy. A total of twelve distinct models were developed by training each of the six YOLO architectures with both RGB and infrared datasets.

TABLE I. Number of Drone Images and Instances for Training, Validation and Testing Datasets

Dataset	**Number of Drone Images**		**Number of Instances**	
	RGB	***Infrared***	***RGB***	***Infrared***
Training	13,869	6,253	13,966	10,251
Validation	1,045	555	1,050	481
Testing	938	252	928	221

D. *Performance Metrics*

Model performance was evaluated using two key metrics: mAP50 (Mean Average Precision at 50% IoU threshold) [28] and frame rate. Intersection over Union (IoU) quantifies detection accuracy by measuring the overlap between predicted and ground truth bounding boxes [25]. mAP50, a standard metric for object detection evaluation [27], represents the mean precision averaged across all classes at 50% IoU threshold. Frame rate, calculated as the inverse of per-image processing time, indicates the model's computational efficiency.

IV. Results and Discussion

A. *Model Training*

As described in Section III, twelve models were developed, with six YOLO variants (YOLOv5n/x, YOLOv10n/x, YOLOv11n/x) trained independently on RGB and infrared datasets. All models demonstrated excellent performance during both training and validation phases, achieving mAP50 scores above 90% with convergence within 100 epochs. Fig. 4 and 5 illustrate the training processes of YOLOv11n models using RGB and infrared drone images, respectively.

TABLE II. Testing Results without Domain Adaptation

Domain		**Model**	**mAP50**	**Number of parameters (M)**	**Frame per second (FPS)**
Train	***Test***				
RGB	RGB	YOLOv5n	0.95	1.76	3.90
IR	IR	YOLOv5n	0.99	1.76	3.84
RGB	RGB	YOLOv5x	0.97	86.17	0.19
IR	IR	YOLOv5x	0.99	86.17	0.18
RGB	RGB	YOLOv10n	0.97	2.69	9.25
IR	IR	YOLOv10n	0.99	2.69	10.15
RGB	RGB	YOLOv10x	0.97	31.59	0.68
IR	IR	YOLOv10x	0.99	31.59	0.69
RGB	RGB	YOLOv11n	0.97	2.58	10.65
IR	IR	YOLOv11n	0.99	2.58	4.74
RGB	RGB	YOLOv11x	0.96	56.83	0.19
IR	IR	YOLOv11x	0.99	56.83	0.84

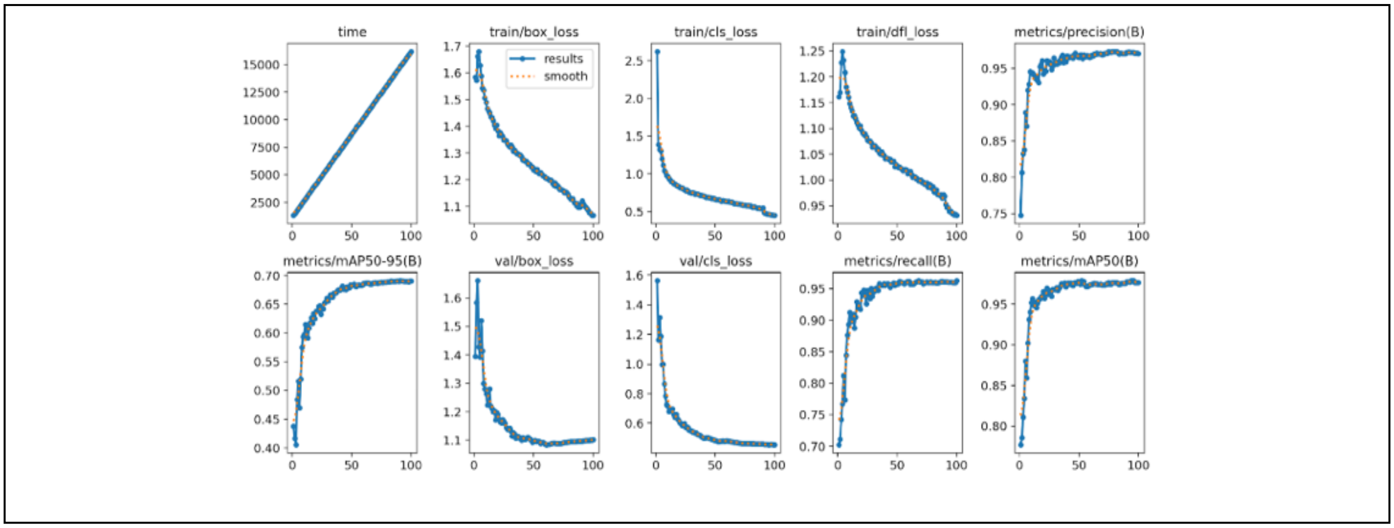

Fig. 4. Training results of the YOLOv11n model using RGB drone images

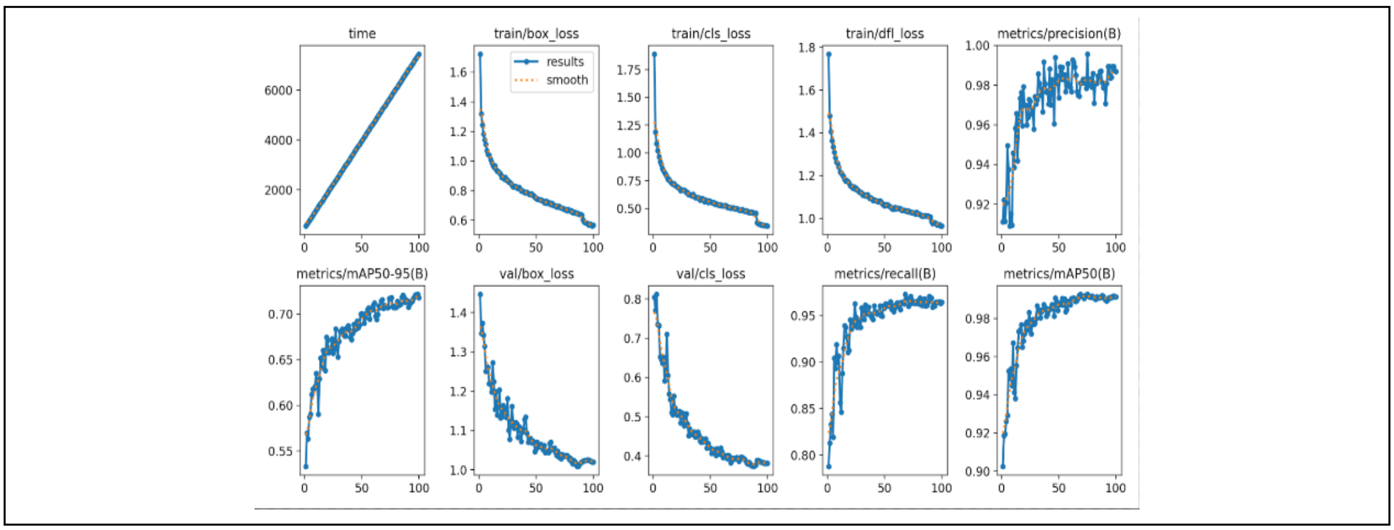

Fig. 5. Training results of the YOLOv11n model using Infrared drone images

B. Model Testing without Domain Adaptation

During the first part of the testing phase, all twelve models were evaluated on a CPU using domain-matched testing datasets. As shown in Table II, the models achieved consistently high mAP50 scores exceeding 95%. However, CPU-based processing yielded relatively low frame rates ranging from 0.18 to 10.65 frames per second (FPS). This processing speed limitation could potentially constrain system performance in scenarios where cameras operate above 11 FPS, despite the models' high detection accuracy.

As evidenced in Table II, all YOLO versions achieved comparable mAP50 values. Among the tested models, YOLOv11n emerges as the most promising candidate for real-time applications, offering the highest frame rate and lowest computational complexity (fewer parameters). These characteristics make it particularly suitable for CPU-based deployments in resource-constrained environments where energy efficiency is critical.

C. Model Testing with Domain Adaptation

The second part of the testing phase evaluated cross-domain performance on a CPU, where models trained on one imaging modality were tested on the other. Models trained on RGB datasets were tested on infrared images, and vice versa. As shown in Table III, RGB-trained models achieved mAP50 scores between 42.6% and 52.4% when detecting drones in infrared images. Fig. 6 illustrates drone detection results on infrared test images using YOLOv11n models: domain-matched training (mAP50: 99%) versus cross-domain RGB training (mAP50: 47.1%). Given that mAP50 scores around 50% are generally considered acceptable in some object detection tasks, the performance of RGB-trained models on infrared images indicates promising cross-domain capabilities. Nevertheless, the CPU-based processing yielded relatively low frame rates ranging from 0.18 to 9.08 frames per second (FPS). This processing speed could potentially limiting system performance in scenarios where cameras operate above 9 FPS, despite the models' promising detection performances.

Conversely, infrared-trained models demonstrated poor generalization to RGB images, yielding mAP50 scores below 1%. Fig. 7 illustrates drone detection results on RGB test images using YOLOv11n models: domain-matched training (mAP50: 97%) versus cross-domain infrared training (mAP50: 0.113%). These results reveal the asymmetric nature of domain adaptation between RGB and infrared modalities in drone detection. This asymmetry likely stems from the relative feature complexity between modalities; RGB images contain richer feature sets that enable some degree of cross-domain detection, while the limited features in infrared images appear insufficient for RGB detection based on the method used in this work.

These findings highlight the potential advantage of leveraging readily available RGB images for infrared drone detection through domain adaptation. Therefore, domain adaptation with CPU-based processing is feasible for drone detection applications. However, further research is needed to enhance both detection performance and processing speed, particularly for applications requiring higher frame rates.

V. Conclusion

This study investigated the feasibility of domain adaptation for drone detection in infrared images using CPU-based YOLO models. We evaluated six YOLO architectures in both Nano (YOLOv5n, YOLOv10n, YOLOv11n) and Extra Large (YOLOv5x, YOLOv10x, YOLOv11x) configurations, training each variant with both RGB and infrared datasets to create twelve distinct models. These models were assessed both with and without domain adaptation. Without domain adaptation, all models achieved excellent performance with mAP50 scores exceeding 95%. With domain adaptation, RGB-trained models demonstrated promising results on infrared images, achieving mAP50 scores between 42.6% and 52.4%. However, infrared-trained models performed poorly on RGB images, yielding mAP50 scores below 1%. This asymmetric performance suggests that the limited features in infrared images are insufficient for RGB detection using our methodology. While detection accuracy showed promise, CPU-based processing achieved relatively low frame rates (0.18-10.65 FPS), potentially limiting applications requiring higher speeds. Our findings demonstrate the feasibility of using RGB-trained models for infrared drone detection through domain adaptation, though further research is needed to enhance both detection performance and processing speed for practical deployment.

Acknowledgment

The authors would like to express sincere gratitude to Wg. Cdr. Peerayudh Saratayon for his idea that initiated this research and to GP. Capt. Chamnan Kumsap, Ph.D., for his support on research writing. Funding support by Defence Technology Institute and Kasetsart University is gratefully acknowledge.

TABLE III. Testing Results with Domain Adaptation

Domain		Model	mAP50	Number of parameters (M)	Frame per second (FPS)
Train	*Test*				
RGB	IR	YOLOv5n	0.484	1.76	3.88
IR	RGB	YOLOv5n	0.00042	1.76	4.07
RGB	IR	YOLOv5x	0.508	86.17	0.18
IR	RGB	YOLOv5x	0.00042	86.17	0.18
RGB	IR	YOLOv10n	0.426	2.69	9.08
IR	RGB	YOLOv10n	0.00042	2.69	10.91
RGB	IR	YOLOv10x	0.482	31.59	0.67
IR	RGB	YOLOv10x	0.000624	31.59	0.66
RGB	IR	YOLOv11n	0.471	2.58	5.74
IR	RGB	YOLOv11n	0.00113	2.58	12.09
RGB	IR	YOLOv11x	0.524	56.82	0.19
IR	RGB	YOLOv11x	0.00042	56.82	0.95

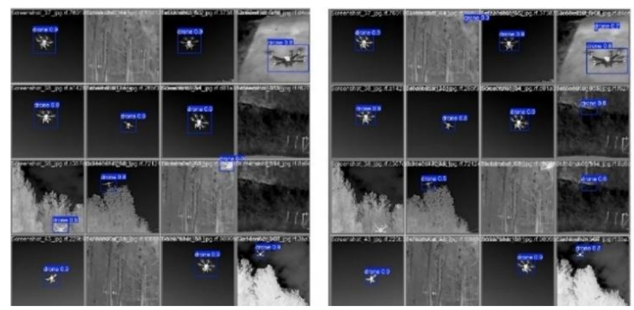

Fig. 6. Drone detection results using YOLOv11n on infrared images: model trained with RGB dataset (left four columns) versus model trained with infrared dataset (right four columns).

Fig. 7. Drone detection results using YOLOv11n on RGB images: model trained with RGB dataset (left four columns) versus model trained with infrared dataset (right four columns).

REFERENCES

[1] Park, S., Kim, H. T., Lee, S., Joo, H., & Kim, H. (2021). Survey on anti-drone systems: Components, designs, and challenges. IEEE access, 9, 42635-42659.

[2] Delleji, T., Slimeni, F., Lafi, M., & Chtourou, Z. (2023, October). Video Data Analytics Dashboard for Anti-drone system. In 2023 International Conference on Cyberworlds (CW) (pp. 134-140). IEEE.

[3] Deshmukh, S., & Sharma, V. (2022, December). An SDR-based anti-drone system with Detection, Traking, Jamming, and Spoofing Capabilities. In 2022 IEEE Microwaves, Antennas, and Propagation onference (MAPCON) (pp. 388-393). IEEE.

[4] Jurn, Y. N., Mahmood, S. A., & Aldhaibani, J. A. (2021, August). Anti-drone system based different technologies: Architecture, threats and challenges. In 2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE) (pp. 114-119). IEEE.

[5] Garcia, A. J., Lee, J. M., & Kim, D. S. (2020, October). Anti-drone system: A visual-based drone detection using neural networks. In 2020 international conference on information and communication technology convergence (ICTC) (pp. 559-561). IEEE.

[6] Farahani, A., Voghoei, S., Rasheed, K., & Arabnia, H. R. (2021). A brief review of domain adaptation. Advances in data science and information engineering: proceedings from ICDATA 2020 and IKE 2020, 877-894.

[7] Bhattacharya, P., Nowak, P., Ahlers, D. K., Holters, M., & Zölzer, U. (2022). CNN based Drone Detection in Infrared Images.

[8] Nageli, V., Kumar, A., Gorthi, R. K. S. S., & Jamal, A. (2023, October). The First Place Solution for ICIP2023 Challenge Infrared Imaging-Based Drone Detection and Tracking in Distorted Surveillance Videos. In 2023 IEEE International Conference on Image Processing Challenges and Workshops (ICIPCW) (pp. 1-5). IEEE.

[9] Bentamou, A., Zein-Eddine, A., Messai, O., & Gavet, Y. (2023, October). Real-Time Drone Detection and Tracking in Distorted Infrared Images. In 2023 IEEE International Conference on Image Processing Challenges and Workshops (ICIPCW) (pp. 3594-3598). IEEE.

[10] Khan, M. U., Misbah, M., Kaleem, Z., Deng, Y., & Jamalipour, A. (2023, June). GAANet: Ghost auto anchor network for detecting varying size drones in dark. In 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring) (pp. 1-5). IEEE.

[11] Zubkov, O., Sheiko, S., Kartashov, V., & Oleynikov, V. (2022, October). Detection of Small Drones in Thermal Infrared Range Using YOLOv5 Neural Networks. In 2022 IEEE 9th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T) (pp. 283-287). IEEE.

[12] Delleji, T., Fekih, H., & Chtourou, Z. (2020, December). Deep Learning-based approach for detection and classification of Micro/Mini drones. In 2020 4th international conference on advanced systems and emergent technologies (IC_ASET) (pp. 332-337). IEEE.

[13] Nalamati, M., Kapoor, A., Saqib, M., Sharma, N., & Blumenstein, M. (2019, September). Drone detection in long-range surveillance videos. In 2019 16th IEEE international conference on advanced video and signal based surveillance (AVSS) (pp. 1-6). IEEE.

[14] Menon, H. P., Vinitha, V., Vishnuraj, K. R., Satheesh, A., & Nikhil, A. P. (2023, December). A Study on YOLOv5 for Drone Detection with Google Colab Training. In 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 1576-1580). IEEE.

[15] Pansare, A., Sabu, N., Kushwaha, H., Srivastava, V., Thakur, N., Jamgaonkar, K., & Faiz, M. Z. (2022, August). Drone Detection using YOLO and SSD A Comparative Study. In 2022 International Conference on Signal and Information Processing (IConSIP) (pp. 1-6). IEEE.

[16] Sony Corporation. (n.d.). *Sony Alpha 6000 user manual*. Retrieved from https://manuals.coolblue.nl/82/sony-alpha-a6000.pdf

[17] FLIR Systems, Inc. (2019). *FLIR Scion thermal monocular user manual*. FLIR Systems, Inc. Retrieved from https://cdn.webshopapp.com/shops/25412/files/302280726/flir-scion.pdf

[18] InfraTec GmbH. (n.d.). *VarioCAM® High Definition user manual*. InfraTec GmbH. Retrieved from https://www.qdindustria.it/wp-content/uploads/2018/12/InfraTec-VarioCAM%C2%AE-High-Definition.pdf

[19] Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 10781-10790). http://doi.ieeecomputersociety.org/10.1109/CVP R42600.2020.01079

[20] Theophanides, T. (2012). Introduction to infrared spectroscopy. Infrared Spectroscopy-Materials Science, Engineering and Technology, 1-10.

[21] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. Machine Learning and Knowledge Extraction, 5(4), 1680-1716.

[22] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[23] Roberts, D. A., Yaida, S., & Hanin, B. (2022). The principles of deep learning theory (Vol. 46). Cambridge, MA, USA: Cambridge University Press.

[24] Amjoud, A. B., & Amrouch, M. (2023). Object detection using deep learning, CNNs and vision transformers: A review. IEEE Access, 11, 35479-35516.

[25] Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 658-666).

[26] Mehlig, B. (2021). Machine learning with neural networks: an introduction for scientists and engineers. Cambridge University Press.

[27] Henderson, P., & Ferrari, V. (2017). End-to-end training of object class detectors for mean average precision. In Computer Vision–ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part V 13 (pp. 198-213). Springer International Publishing.

[28] Kukil. (2022, August 9). *Mean Average Precision (mAP) in Object Detection*. LearnOpenCV. https://learnopencv.com/mean-average-precision-map-object-detection-model-evaluation-metric/

[29] BigData. (2023, December). Drone Detect Dataset. Roboflow Universe. Roboflow. Retrieved August 11, 2024, from https://universe.roboflow.com/bigdata-esfic/drone-detect-t7o67

[30] customyolo. (2023, November). Drones Dataset. Roboflow Universe. Roboflow. Retrieved July 24, 2024, from https://universe.roboflow.com/customyolo-k3lvg/drones-rom1a

[31] Kalra, A., Stoppi, G., Brown, B., Agarwal, R., & Kadambi, A. (2021). Towards rotation invariance in object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 3530-3540).