Defense of Military Installations from UAV-borne attacks using Deep Learning

Nicholas Schaefer
Department of International Studies and Political Science,
Virginia Military Institute, Lexington, Virginia, USA.
schaeferng23@mail.vmi.edu

Jafaar M. Alghazo
Department of Electrical and Computer Engineering,
Virginia Military Institute, Lexington, Virginia, USA.
alghazojm@vmi.edu

Sherif E. Abdelhamid
Department of Computer and Information Sciences,
Virginia Military Institue,
Lexington, Virginia, USA.
abdelhamidse@vmi.edu

Ghazanfar Latif
Department of Computer Science,
Prince Mohammad Bin Fahd University,
Alkhobar, Saudi Arabia.
glatif@pmu.edu.sa

***Abstract*— Consumer-grade Unmanned Aerial Vehicles (UAVs) are becoming more common capabilities on the modern battlefield, finding use by both formal standing armies and non-state sponsored organizations with small budgets. The threats posed by these UAVs are varied, ranging from intelligence gathered from reconnaissance, spotting for indirect fires, or attacking with a payload on the UAV itself. These dangers create new challenges that militaries must adapt to ensure soldiers are protected and mission completion is possible despite the threat of UAV interdiction. In this paper, we propose an AI object detection model that is capable of identifying UAVs in the visible spectrums and distinguishing them from images that contain no UAVs. This model can be used as a targeting system for anti-UAV countermeasures. The dataset used to train the model consists of 2,000 images. 1000 images are of UAVs and 1000 contain no UAVs. The models we tested were ResNet18, ResNet50, and GoogleNet. GoogleNet achieved the best results, yielding a precision of 0.995, recall of 0.995, F1-score of 0.995, and test accuracy of 98.44%. These results and the initial dataset present a good base for researchers to explore and design a practical solution for defense against small drone attacks.**

Keywords—Convolutional Neural Network, ResNet, GoogleNet, AlexNet, Transfer Learning, Drones, UAV, military.

I. INTRODUCTION

Artificial Intelligence (AI) has found many important roles in our world. It helps conduct studies, ensures security, and even powers some of the more mundane aspects of everyday life. It is also used in modern warfare in a variety of roles. Most prominently, it has been integrated into various platforms targeting systems for the identification of threats and electing how to counter or destroy them, which assists in the automation of the protection of military forces. As threats change, these models must be updated or replaced to ensure they can continue providing effective security.

Another emerging technology that has found itself at home in war is the Unmanned Aerial Vehicle (UAV). UAVs pose significant benefits to the parties that use them. They can be used without exposing any human life to risk. They can come in all shapes and sizes, which their capabilities are tied to. However, perhaps most importantly, they are easily proliferated and modified into a hard-to-detect vehicle that can launch a new attack on one's enemy. Of course, these same advantages translate to disadvantages for anyone on the side opposing any party that uses UAVs. Military, law enforcement, and civilians can all be subject to the problems posed by UAV threats from any number of actors, including state governments, terrorist organizations, or other negligent or malicious civilians.

UAVs offer many beneficial and benign uses for users. These include agricultural, environmental monitoring, delivery, internet service, maritime monitoring, and unexploded ordinance detection and disposal [1]. These kinds of services are very important to society for various reasons, from the preservation of our ecosystems to logistics, resource acquisition, safety, and security. However, UAVs are finding a large number of military applications as well. UAVs were most notably used by the US during the war on terror, but the common perception of these UAVs is of the larger airframes flying many kilometers high hitting targets with Hellfire missiles. The foil of this is the types of UAVs being used by groups with less funding and infrastructure. These are most often Consumer Off the Shelf (COTS) UAVs that have been refitted with various payloads (some deadly) and repurposed into more deadly tools [2]. This can be observed quite frequently in the current war in Ukraine. Russian and Ukrainian forces use UAVs to drop grenades and small explosives, carry suicide payloads, conduct reconnaissance, and spot indirect fires [3]. Sometimes a single drone is used to reduce the signature of an activity. Others, swarms of dozens are used to make any activities harder to stop. They have found a role that has fundamentally changed the way people fight.

The motivation for researching how AI may be used to reduce the threats and problems created by modern UAV tactics is to help ensure the safety of military personnel, assets, and installations. As previously stated, military forces are becoming increasingly vulnerable to COTS UAV attacks and reconnaissance. ML models are capable of detecting UAVs in images and can therefore be adapted to this use. Such a model can be used as the basis for a targeting system for an asset designed to shoot down UAVs or counter them with other means. Key points in military installations can be protected and forces in the vicinity of vehicle-mounted systems can also enjoy the security it provides.

The aim of this paper is to explore the use of deep learning algorithms for the detection of UAVs. We also aim to develop a dataset for our purpose because we could not find a suitable one

979-8-3503-2759-5/23/$31.00 ©2023 IEEE
DOI 10.1109/CSCE60160.2023.00047

for our use. produce an AI-powered model that is capable of detecting UAVs through camera sensors. It must be highly accurate to avoid targeting small objects that are not UAVs. It needs to be a very quick model, as there are scenarios where UAVs are suddenly present and the vulnerable parties may only have seconds to react to the threat. It will play an extremely critical role in ensuring the safety of people in our world today.

The rest of the paper is organized in the following way: section 2 reviews current literature on the topic being researched. Section 3 discusses the proposed method that will be used and how it was trained, including information on the selected dataset. Section 4 covers test results and analysis, and section 5 concludes the paper by highlighting the conclusion and future work. Section 6 lists all references used in this paper.

II. Literature Review

Due to the extremely rapid transition to the use of consumer UAVs as spotters and payload carriers in combat by groups that are required to be resourceful, especially in Ukraine, the scientific community is now looking for new ways to combat this developing use. There are many soldiers' and civilians' lives at risk due to this new way to wage war, and developing an AI model with the capability of detecting and identifying UAVs would lead to a new capability that could substantially reduce bloodshed and tragedy.

In [4], the authors proposed a method that uses an enhanced You Only Look Once (YOLO) v5 deep learning model to detect and identify different Unmanned Aerial Vehicles and their payloads. They named this model YOLOv5s(PANet), with PANet standing for Path Aggregation Network. Their dataset is manually generated and consists of 5,460 images of drones in various environments and 1,709 images of drones with various attached payloads. The drones that appear in the dataset include Anafi Extended, DJIFPV, Mavic2 Enterprise Dual, MAVIC Air, and EFT-E410S. The payloads include various objects such as missiles, bombs, cameras, mounted guns, and packages. They reserved 70% of the data for training, 20% for testing, and 10% for validation. They report that all the tested models, which include their proposed model, YOLOv5m, and YOLOv5s, achieved a mean average precision (mAP) and recall of 100% when detecting UAVs in cloudy environments. They found that YOLOv5(PANet) yields a greater mAP of 99% compared to the other two model's mAPs of 97% and 98% when detecting small UAVs in evening environments. When testing payload detection on YOLOv5s and YOLOv5s(PANet), the authors found that, although their model did outperform YOLOv5s, neither model performed that well. YOLOv5(PANet) yielded mAPs ranging from 50% to 95% and recalls of 47.1% to 75.8% depending on the type of payload being detected. They state this is a result of some of the payloads being hidden inside the UAV and not being properly captured in the image. YOLOv5s(PANet) received an F1-score of 89.5%.

In [5], the authors propose a possible low-cost AI-powered counter-drone system that can detect, identify, and eliminate hostile UAVs to alleviate the currently problematically expensive options available presently. Their solution is to use a detection algorithm on board a UAV that will track and intercept hostile UAVs. They use a total of three models. Model 1 has a Convolutional Neural Network (CNN) architecture. Model 2 uses YOLOv3 with pre-trained weights from Darknet-53. Model 3 uses an optimized version of the EfficientNet-B0 algorithm. The dataset they used to train, test, and validate the models was created using Airsim Simulator, which uses Unreal Engine to generate realistic scenarios, to generate 2,000 images of drones in random positions in the generated environment. They report Models 1, 2, and 3 have mAPs of 83.63%, 84.93%, and 89.77% respectively. Models 1, 2, and 3 achieved accuracies of 85%, 91%, and 86%, and F1-scores of .90, .94, and .91 respectively. One limitation of this study is that approximately 10% of auto-labeled images were inaccurate and had to be removed from the dataset, which was caused by the hunting and hostile UAVs being too close together, which results in extremely high relative speed.

In [6], the authors proposed the use of counter-UAVs to detect intruding UAVs, which is more efficient than human intervention. They use an algorithm called Deep Q-Learning from Demonstrations (DQfD) to detect flying UAVs and capture them. They train DQfD using Reinforced Learning (RL) with incremental rewards and trials of detecting and capturing UAVs on Airsim. The counter-UAV could take one of five actions, which were moving forward, yawing left or right, and ascending or descending. They tested nine different deep-reinforcement learning algorithms to see which was most effective. Simulations were rendered by an NVIDIA GeForce RTX 3060 Ti with 8GB VRAM. They report that their lowest resulting success rate was 5% and their highest was 98%. The 5% success rate was caused by the model routinely crashing the UAV. In [7], the authors proposed a new object detection dataset intended to train visually based object detection machine learning algorithms to detect multiple Unmanned Aerial Vehicles (UAVs) using a camera. It was constructed to create a large dataset of UAVs specifically tailored to object detection, which the author claims had not been done prior to the creation of their dataset. This dataset is made of real-world hand-labeled images of drones intended for the visual detection of UAVs. It consists of 51,446 images, some of which were manually labeled and some of which were labeled by an Artificial Neural Network (ANN) based model used to make the process semi-automated. The dataset was generated with the goal of sampling a broad range of environmental variables, such as time of day, weather, landscape, and UAV position relative to the camera. They used this data to train 819 instances of various Convolutional Neural Networks (CNN) and 603 Haar Cascades. The Haar Cascades resulted in more easily deployed but less effective models, yielding a maximum accuracy of 55.4%, precision of 81.7%, recall of 18.1%, and F1-score of 28.6%. In contrast, the CNN-based models were not as easily deployed but had much better results, including a maximum precision of 89.3%, recall of 49.3%, specificity of 95.7%, an accuracy of 70.3%, and F1-score of 62.7%They find that HAAR Cascades are easily deployed and perform moderately, achieving a modest 55% accuracy and 0.32 F1 score, but pale in comparison to more comprehensive Deep Neural Networks which require more computing power and achieved accuracies of 70% and an F1 score of 60.2.

In [8], the authors proposed the use of the YOLOv5 algorithm in visual UAV detection systems. They used a dataset called Det-Fly, which is comprised of 13,271 hand-labeled images of DJI-Maveric drones against various simple and complex backgrounds. They trained both YOLOv5 and Faster R-CNN with this data. They then tested these two algorithms against each other and found that Faster R-CNN was prone to false positives, as well as false positives in simple scenarios, while YOLOv5 performed generally better in both simple and complex scenarios. In [9], the authors proposed using a multi-agent depth deterministic policy gradient (MADDPG) algorithm to control swarms of drones designed to counter hostile drone swarms. They used 50,000 rounds of simulated training to train their AI through unsupervised learning and found that there was a problem causing friendly UAVs to overshoot and miss their targets. After implementing a rule to mitigate this, they conducted 500 test rounds and found that their success rate had jumped from 63% to 81%.

In [10], the authors proposed an Unmanned Aerial Vehicle (UAV) detection system that uses machine learning and object detection to alleviate various problems with radar detection systems. They used AI to control surveillance UAVs, which searched for and identified potentially hostile UAVs in video frames. They elected to use the Haar-Feature-based Cascade Classifier trained on 7000 positive images (including images modified using distortion) and 3019 negative images gathered from Google and http://face.urtho.net/. They reported a maximum accuracy of 91.6% and an average of approximately 89%. In [11], the authors proposed the use of the NVIDIA Jetson TX2 as an object detector to detect unwelcome UAVs. They used YOLOv3 with pre-trained weights and transfer learning to train the algorithm to detect UAVs. They manually sorted 1435 images and used various forms of data augmentation to expand the dataset to 7175 images. They reported an average confidence level of 88.9% resulting from their tests. They also reported that a limitation of their study was the low processing power of the Jetson TX2, which resulted in low framerates.

In [12] the authors proposed a real-time UAV detection algorithm based on the necessity of real-time detection to combat unauthorized UAV activity and make distinctions between UAVs and birds. They used MobileNetV2 CNN and obtained a dataset of 24,075 images including 10,155 images of UAVs, 4572 images of birds, and 9348 background images. They reported an accuracy of 99.83% and reported a limitation on the ability of their algorithm to detect UAVs moving near other small moving objects. In [13], the authors proposed a UAV detection algorithm to replace radar detection systems, which have trouble detecting and making distinctions between small objects. They used a YOLOv3 object detector and a multiclass dataset consisting of more than 10,000 images of tricopter, quadcopter, and hexacopter UAVs. They reported training resulted in approximately 95% accuracy and approximately 98% recall. To remedy the possibility of the model incorrectly identifying the type of UAV it detects, they suggest RF signal or audio detection be used in conjunction with their model.

In [14], the authors proposed the use of YOLOv4 for automated visual UAV detection to correctly distinguish between UAVs and birds to defend against the misuse of UAVs. They collected images from Kaggle and Google to build their dataset of 2,395 images, which included 479 bird images and 1,916 images of UAVs. They reported testing the model on detecting DJI Mavic Pro and DJI Phantom III drones resulted in an mAP of 74.36%, precision of 0.95, recall of 0.68, and an F1 score of 0.79. They also reported that the altitude of the object being detected created a limitation for the model because higher-altitude objects were more difficult to detect and correctly identify. AI-driven models are already being widely used detection and classification of various objects in images. In [15], the authors proposed an AI model to detect and track people using UAVs fitted with cameras in order to ensure the efficiency and reliability of the tasks they perform. They used Single Shot Detector (SSD) trained on 5,100 images in a binary dataset that either contained a person or no person. They then put their model on board a Parrot AR Drone 2 which would represent a UAV tasked to detect people. They reported a sensitivity of 0.98 and precision of 0.99, which were superior to results achieved by a YOLO object detection algorithm.

In [16], the authors proposed an algorithm that rapidly detects UAVs to protect various security risks from being exploited by malicious UAV users. Their detection algorithm is based off of the improved CenterNet, which created a feature map and identifies UAVs by their features. They trained their model on a dataset consisting of 1,800 images of various types of UAVs, reserving 1,300 images for training and 500 for testing. They tested their proposed model against Faster-RCNN, YOLOv4, and CenterNet(dla-34) and found that, although their model had a lower average precision (0.930 compared to YOLOv4's 0.971), their model was able to maintain an FPS of 143.32, which meant the model could detect drones more quickly and make repeated assessments faster. Future implementation of this model would be to enable steps to counter malicious UAVs, such as alarming or jamming. In [17], the authors proposed a drone detection model based on YOLOv4 to secure no-fly zones in order to protect against economic losses and passenger safety. They trained and tested YOLOv4, YOLOv3, and SSD to determine which model was best at detecting UAVs. Their dataset contained 1,540 images of DJI-Phantom, DJI-Inspire, and XIRO-Xplorer UAVs and 556 images of these same UAVs, which were rotated and flipped to increase the amount of data. The final size of the dataset is 3,218 images. In testing, they found that YOLOv4 was the overall best model with an accuracy of 89.32% and recall of 92.48%, which outperformed both YOLOv3 and SSD's results.

In [18], the authors proposed a model that would be able to detect and identify UAVs in order to counter the increasing proliferation of UAVs for malicious purposes. They focus specifically on a solution to accurately detect and identify miniature UAVs at long distances in real-time. They used the PANet algorithm as a backbone for their model, which they tested against other detection algorithms to evaluate its effectiveness. They used two datasets. One consists of 7,200 images of six different types of UAVs, such as Anafi Extended, DJI-Phantom, DJI-FPV, EFTE410S, Mavic-Ent, and Mavic-Air at various altitudes and environments. The second dataset is focused on payload identification and includes 3,600 images of medical supplies, spy cameras, sealed packages, containers, guns, food items, missiles, and explosives. The authors report their model achieved an average mAP of 82.75%, sensitivity of 87.97%, specificity of 47.1%, G-mean of 63.38%, and an F1-score of 82.75%.

In [19], the authors proposed the use of transfer learning to train a model to detect and classify UAVs to prevent the malicious use of UAVs. They test multiple kinds of transfer learning classification algorithms against each other, including Inception V3 and ResNet 101. Their dataset included multiple images of types of UAVs, birds, kites, and landscapes. The categories of both UAV positive and negative images are 10,080. They found that Inception V3 yielded superior results with an accuracy of 96.82%, precision of 96.75%, and recall of 96.98%. In [20], the authors proposed a model that is capable of detecting and identifying swarms of miniature UAVs in order to combat terrorist and military UAV activity. Their model was a modified YOLOv5 object detection algorithm, which they tested against Visual Geometry Group 16 (VGG-16), GoogleNet, and MobileNet. Their custom dataset consisted of images of DJI-Phantom, Mavic-Air, and Mavic enterprise drones in various conditions, including cloudy, sunny, and evening environments. They added images of birds downloaded from Kaggle to improve their dataset and reach a total of 1,000 images. They split this data into the 70:20:10 ratio for training, testing, and validation respectively. They found that their proposed YOLOv5 model achieved the best results, including a precision of 94.30%, recall of 100%, and F1-score of 97%. These results were superior to those produced by the other models tested.

More literature demonstrates how AI has been used for visual detection in other fields. In [21], the authors proposed using an AI model to detect and classify types of brain tumors to expedite diagnosis, increase accuracy, and reduce the cost of such a procedure. They used a Support Vector Machine (SVM) with Rectified Linear Unit (ReLU) activation function trained on the MICCAI BraTS 2018 dataset, which consists of 40,300 images. 24,180 of these were High-Grade Glioma cases, 16,120 were Low-Grade Glioma Cases, and 4,250 had no tumors. They tested SVM against Multilayer Perception (MLP), Random Forest (RF), and Naïve Bayes (NB), and found SVM had superior results, producing a maximum accuracy of 96.19%, precision of .958, recall of .851, and an F1-score of 0.870. In [22], the authors proposed a deep-learning hybrid framework that is able to detect objects to aid the navigation of self-driving cars. They use a YOLOv4 model trained on the Berkeley DeepDrive 100k (BDD100k) dataset, which consists of 70,000 images containing various objects with labels reflecting a multiclass system, including pedestrians, traffic lights, and other vehicles. Their model was tested against Single-Shot Detector (SSD), Wasserstein Loss-based Model for Object Detection (WLOD), and unmodified YOLOv3 and YOLOv4. They conducted their experiments on an Ubuntu machine powered by an Intel Core i7-5930K, which yielded an mAP of 52.7, which is 2.6 greater than the next highest-performing model tested, which is YOLOv4. One limitation of this model is that it struggles to perform as well in inclement weather.

In [23], the authors proposed an AI-enabled detection system for use aboard UAVs for the purpose of detecting traffic objects and predicting pedestrian behavior in order to mitigate traffic accidents. They use a Feature Fusion and Scaling-based SSD (FS-SSD) network trained by the Car Parking Lot (CARPK) dataset and Stanford Done Dataset (SDD), which consisted of 69,673 multiclass images including captures of pedestrians, bikers, and cars. They trained their FS-SSD model against many other models, including Faster R-CNN, YOLOv3, and multiple variations of SSD. They report their FS-SSD model achieved a maximum mAP of 66.42%, which outperformed the other models by at least 4%. One limitation of this study is that, despite being the most effective model tested, the model seems to trade Frames Per Second (FPS) to achieve superior results, which places a limitation on performance.

III. Dataset and Methodologies

A. Dataset Description

The dataset used in this study contains 224x224x3 images of UAVs conducting various maneuvers and from different angles and clear sky with landscape. The UAVs in the images are many different types of COTS UAVs. 2,000 images make up this dataset, in which an AI-powered model was trained, tested, and validated. There are 1,000 images of UAVs and 1000 images containing no UAVs. This dataset was split into the 70%-20%-10% ratio for training, testing, and validation respectively. The images form a multiclass dataset, the labels being either "Drone" or "No Drone." The images with UAVs have skies as the background and the images without UAVs are of clear skies, some of which have some foreground elements and landscape. All of these images are in grayscale. Figure 1 includes some sample images from the dataset to demonstrate what the images generally look like.

Fig. 1. Four sample images from the dataset. The two on the left contain UAVs, the two on the right contain no UAVs and only clear skies and some landscape.

B. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are artificial intelligence models that are inspired by the way the human brain functions. It uses many layers of connected nodes, which are called neurons, with different weights and parameters to generate outputs in response to different inputs. These networks are a subset of Artificial Neural Networks (ANNs), which consist of at least three layers. As shown in Figure 2, these are the input layer, the hidden layer(s), and the output layer. Although three layers are the minimum, there is no upward limit on the number of layers a CNN can have (aside from technological limits). As these networks are trained, the model self-optimizes itself by adjusting the weights of each neuron in the network [24].

CNNs were designed specifically for the purpose of image processing and detection of patterns or objects using pixel values as inputs. The discernable factor that makes a CNN distinct is the specific layer types in addition to input layers. These are convolutional layers, pooling layers, and fully-connected layers. The convolutional layer uses the weights at each node to determine outputs. Pooling layers perform down sampling to reduce the number of parameters within the activation. Fully-connected layers produce classification scores based on outputs [24].

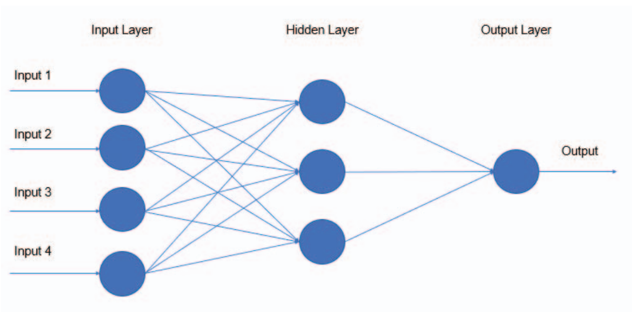

Fig 2. An example of a minimalistic three-layered neural network.

C. Transfer Learning

Two common hurdles to AI model development are the time and amount of data required to sufficiently train a model. The time required may not fit the timeline of researchers or there may not be enough data readily available to train and test the model. One common solution is transfer learning. As shown in Figure 3, transfer learning is the concept of expediting the development of a solution to one problem by using knowledge gained from the solution of another problem. In the field of AI, transfer learning is the practice of taking a model generated during another study and using it to prototype a new model [25].

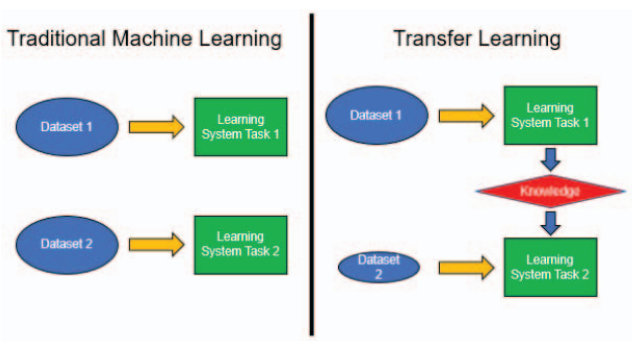

Fig 3. A diagram illustrating the difference between traditional machine learning and transfer learning.

There are six types of transfer learning. Domain Adaptation is where the marginal probability distributions are different, but features are similar between the target and source. Domain Confusion is when models learn domain-invariant features, which improves transferability. Multitask learning is when the model learns both target and source tasks are learned simultaneously. One-Shot Learning is when a model tries to generate an accurate output with very little exposure to data. Zero-Shot learning is when a model learns a new concept by being exposed to unlabeled data it has never seen before. Meta-Learning is the of models "learning to learn" and using only prior knowledge to optimize themselves at performing new tasks [25].

D. Algorithms

In this project, we explore the following deep-learning algorithms.

GoogleNet is an inception-based ML model used for object detection and identification in images. It was introduced in 2014 by Google and features 22 layers in its neural network. The aspect of GoogleNet' S relatively unique inception model is its one-dimensional series configuration. As shown in Figure 4, nine of these inception models make up the GoogleNet architecture, which dictates the output of the model [26].

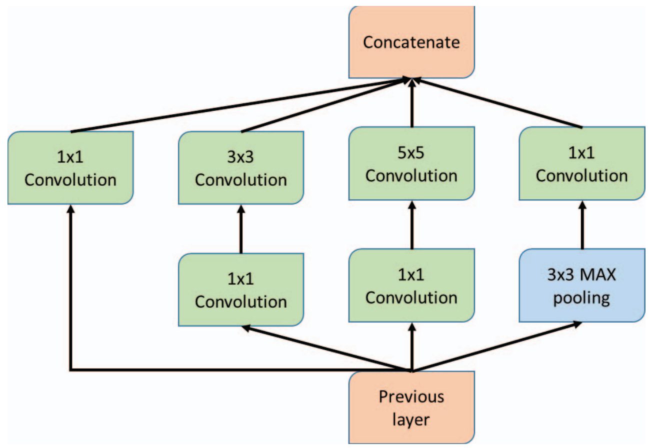

Fig 4. GoogleNet Architecture [26].

AlexNet is an ML model used for object detection in images. It was introduced in 2012 by Alex Krizhevsky et al. It is comprised of 8 layers, which are shown in Figure 5. The first five are convolutional and the remaining are all fully connected. It uses a ReLU activation function, which is applied to every fully connected and convolutional layer in the model. The output layer traditionally has 1,000 nodes for the classification of objects [26].

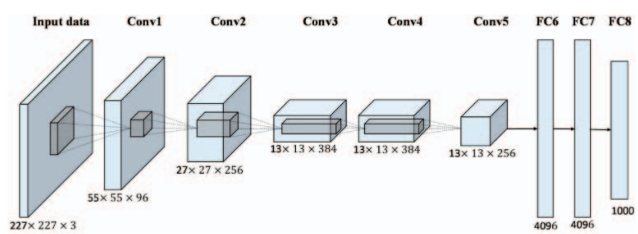

Fig 5. AlexNet architecture [26].

ResNet 18 and ResNet50 are Convolutional Neural Networks (CNNs) created to solve the vanishing gradient problem. This problem occurs in deep networks and causes the loss function to calculate gradients as zero, resulting in the values of weights never changing. In ResNet models, gradients flow through skip connections backward from lateral layers to initial filters. As shown in Figure 6, each layer is made up of many blocks, each with its number of operations. The primary difference between ResNet 18 and Resnet 50 is the number of layers, which have 18 and 50 layers respectively [27].

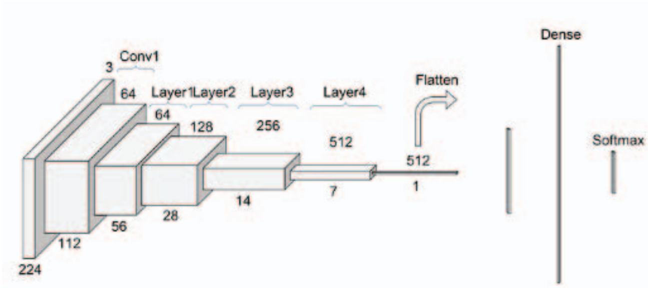

Fig 6. Basic ResNet architecture [27].

VGG16 and VGG55 are the two other CNNs selected for this project. VGG (Visual Geometry Group) utilizes blocks consisting of 2D convolution and Max Pooling layers. VGG16 is an improvement of AlexNet and uses 3x3 filters instead of AlexNet's larger filters, which are the smallest size possible to capture vertical and lateral information. It has a simple architecture but contains a robust 138 million parameters. The architecture consists of an input layer, convolutional layers, ReLU activation, hidden layers, pooling layers, and fully connected layers., which are shown in Figure 7. As with ResNet, VGG16, and VGG55 have 16 and 55 layers respectively [28].

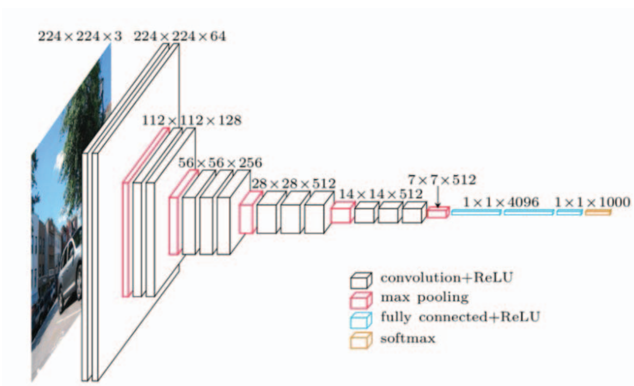

Fig 7. Basic VGG Architecture [28].

IV. Results

This experiment was set up for 10 epochs per ML model. GoogleNet achieved the highest precision, recall, and F1 score, all of which was 99.5%. It also achieved a test accuracy of 98.44%. ResNet18 was the second best, achieving a precision, recall, and F1-score of 98.5% and a test accuracy of 97.66%. ResNet50 was the third best, achieving a precision of 97.6%, a recall, and F1-score of 97.5%, and a test accuracy of 97.6%. Table I displays the results of the experiment.

Table I. Comparison of the results of the experiment of using GoogleNet, ResNet 18, and ResNet 50 transfer learning models to detect UAVs.

Model	Precision	Recall	F1-Score	Test Loss	Test Accuracy
ResNet18	0.985	0.985	0.985	0.065	95.66
GoogleNet	0.995	0.995	0.995	0.061	98.44
ResNet50	0.976	0.975	0.975	0.061	97.66

Figures 8 and 9 illustrate validation learning curves for accuracy and loss using GoogleNet. The curve demonstrates an ideal saturation point by showing accuracies approaching 100% and approaching a loss of 0. This reinforces that our model does not suffer from any overfitting or underfitting issues.

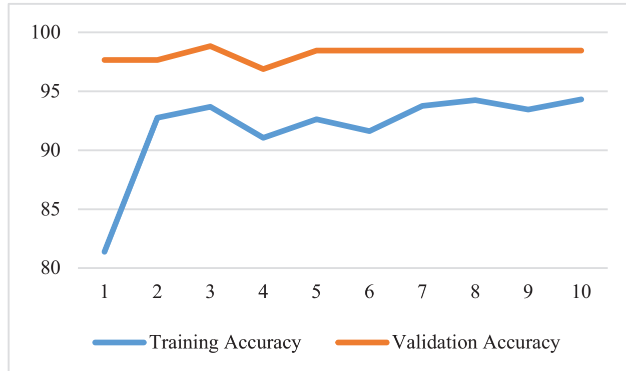

Fig 8. GoogleNet transfer learning-based training/validation learning curves depicting accuracy.

Fig 9. GoogleNet transfer learning-based training/validation learning curves depicting loss.

Figures 10 and 11 illustrate validation learning curves for accuracy and loss using ResNet18. The curve demonstrates an ideal saturation point by showing accuracies approaching 100% and approaching a loss of 0. This reinforces that our model does not suffer from any overfitting or underfitting issues.

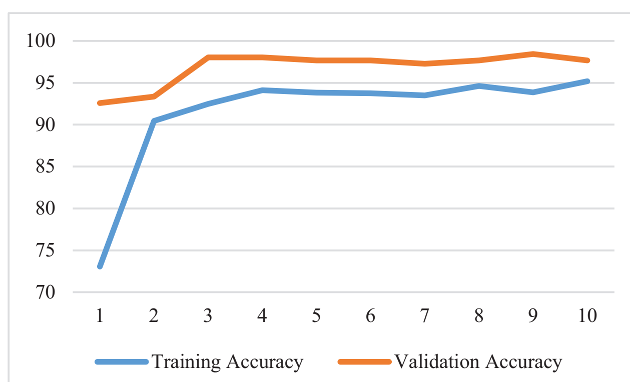

Fig 10. ResNet18 transfer learning-based training/validation learning curves depicting accuracy.

Fig 11. ResNet18 transfer learning-based training/validation learning curves depicting loss.

Figures 12 and 13 illustrate validation learning curves for accuracy and loss using ResNet50. The curve demonstrates an ideal saturation point by showing accuracies approaching 100% and approaching a loss of 0. This reinforces that our model does not suffer from any overfitting or underfitting issues.

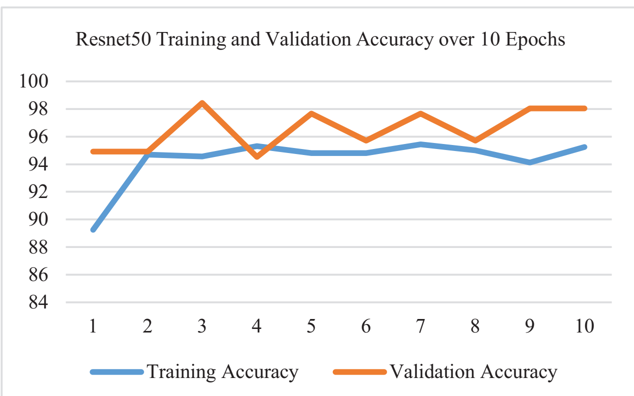

Fig 12. ResNet50 transfer learning-based training/validation learning curves depicting accuracy.

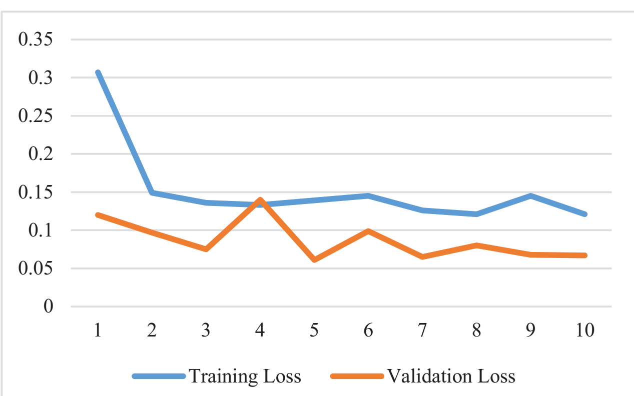

Fig 13. ResNet50 transfer learning-based training/validation learning curves depicting accuracy.

V. Conclusion

Consumer-grade Unmanned Aerial Vehicles (UAVs) are becoming more common capabilities on the modern battlefield, finding use by both formal standing armies and non-state sponsored organizations with small budgets. The threats posed by these UAVs create new challenges that militaries must adapt to ensure soldiers are protected and mission completion is possible despite the threat of UAV interdiction. In this paper, we propose an AI object detection model that is capable of identifying UAVs in both the visible spectrums and distinguishing them from clear sky images with landscape. This model can be used as a targeting system for anti-UAV countermeasures. A newly developed dataset was used that has 1000 images containing UAVs and 1000 images of clear sky with no UAVs.. The models we tested were ResNet18, ResNet50, and GoogleNet. GoogleNet achieved the best results, yielding a precision of 0.995, recall of 0.995, F1-score of 0.995, and test accuracy of 98.44%. These results and the initial dataset present a good base for researchers to explore and design a practical solution for defense against small drone attacks.

The information gathered from this project can be used as a start for more research to develop counter-UAV systems that can be deployed on military installations and vehicles. The dangers to military personnel posed by hostile UAVs are broad, ranging from enemy intelligence gathering to enemy attacks on military assets via UAV-mounted munitions. By using the models tested or a successor to them, a targeting system can be developed for a capability that can spot, identify, and possibly even destroy hostile UAVs before they can inflict any harm.

Acknowledgments

I would like to thank the Virginia Military Institute honors program for making this class possible. I would also like to thank Dr. Abdelhamid and Dr. Latif for their contributions and flexibility as we worked through this. Finally, I want to express my deepest thanks to LTC Alghazo for mentoring me through the process and never giving up on me as I navigated this discipline for the first time.

Funding: This work was supported in part by the Commonwealth Cyber Initiative, an investment in the advancement of cyber R&D, innovation, and workforce development. For more information about CCI, visit cyberinitiatives.org.

References

[1] Muchiri, G. N., & Kimathi, S. (2022, April). A review of applications and potential applications of UAV. In Proceedings of the Sustainable Research and Innovation Conference (pp. 280-283).

[2] Hartmann, K., & Giles, K. (2016, May). UAV exploitation: A new domain for cyber power. In 2016 8th International Conference on Cyber Conflict (CyCon) (pp. 205-221). IEEE.

[3] Russell, S. (2023). AI weapons: Russia's war in Ukraine shows why the world must enact a ban. Nature, 614(7949), 620-623.

[4] Ajakwe, S. O., Ihekoronye, V. U., Akter, R., Kim, D. S., & Lee, J. M. (2022, January). Adaptive drone identification and neutralization scheme for real-time military tactical operations. In 2022 International Conference on Information Networking (ICOIN) (pp. 380-384). IEEE.

[5] Cetin, E., Barrado, C., & Pastor, E. (2021). Improving real-time drone detection for counter-drone systems. The Aeronautical Journal, 125(1292), 1871-1896.

[6] Çetin, E., Barrado, C., & Pastor, E. (2022). Countering a Drone in a 3D Space: Analyzing Deep Reinforcement Learning Methods. Sensors, 22(22), 8863.

[7] Pawełczyk, M., & Wojtyra, M. (2020). Real world object detection dataset for quadcopter unmanned aerial vehicle detection. IEEE Access, 8, 174394-174409.

[8] Samaras, S., Diamantidou, E., Ataloglou, D., Sakellariou, N., Vafeiadis, A., Magoulianitis, V., ... & Tzovaras, D. (2019). Deep learning on multi sensor data for counter UAV

[9] Xiang, L., & Xie, T. (2020, December). Research on UAV swarm confrontation task based on MADDPG algorithm. In 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE) (pp. 1513-1518). IEEE.

[10] Lee, D., La, W. G., & Kim, H. (2018, October). Drone detection and identification system using artificial intelligence. In 2018 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 1131-1133). IEEE.

[11] Xun, D. T. W., Lim, Y. L., & Srigrarom, S. (2021, January). Drone detection using YOLOv3 with transfer learning on NVIDIA Jetson TX2. In 2021 Second International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP) (pp. 1-6). IEEE.

[12] Seidaliyeva, U., Akhmetov, D., Ilipbayeva, L., & Matson, E. T. (2020). Real-time and accurate drone detection in a video with a static background. Sensors, 20(14), 3856.

[13] Behera, D. K., & Raj, A. B. (2020, May). Drone detection and classification using deep learning. In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1012-1016). IEEE.

[14] Singha, S., & Aydin, B. (2021). Automated drone detection using YOLOv4. Drones, 5(3), 95.

[15] Rohan, A., Rabah, M., & Kim, S. H. (2019). Convolutional neural network-based real-time object detection and tracking for parrot AR drone 2. IEEE access, 7, 69575-69584.

[16] Latif, G., Ben Brahim, G., Iskandar, D. A., Bashar, A., & Alghazo, J. (2022). Glioma Tumors' classification using deep-neural-network-based features with SVM classifier. Diagnostics, 12(4), 1018.

[17] Li, Y., Wang, H., Dang, L. M., Nguyen, T. N., Han, D., Lee, A., ... & Moon, H. (2020). A deep learning-based hybrid framework for object detection and recognition in autonomous driving. IEEE Access, 8, 194228-194239.

[18] Ajakwe, S. O., Ihekoronye, V. U., Kim, D. S., & Lee, J. M. (2022, September). Tractable Minacious Drones Aerial Recognition and Safe-Channel Neutralization Scheme for Mission Critical Operations. In 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA) (pp. 1-8). IEEE.

[19] Meng, W., & Tia, M. (2020, December). Unmanned Aerial Vehicle Classification and Detection Based on Deep Transfer Learning. In 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) (pp. 280-285). IEEE.

[20] Ihekoronye, V. U., Ajakwe, S. O., Kim, D. S., & Lee, J. M. (2022). Lightweight CNN Model For Real-Time Recognition of Miniaturize Fleet of UAVs. 한국통신학회 학술대회논문집, 203-206.

[21] Zhizhong, X., Jingen, W., Zhenghao, H., & Yuhui, S. (2020, October). Research on multi UAV target detection algorithm in the air based on improved CenterNet. In 2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE) (pp. 367-372). IEEE.

[22] Shi, Q., & Li, J. (2020, October). Objects detection of UAV for anti-UAV based on YOLOv4. In 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT (pp. 1048-1052). IEEE.

[23] Liang, X., Zhang, J., Zhuo, L., Li, Y., & Tian, Q. (2019). Small object detection in unmanned aerial vehicle images using feature fusion and scaling-based single shot detector with spatial context analysis. IEEE Transactions on Circuits and Systems for Video Technology, 30(6), 1758-1770.

[24] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.

[25] Ribani, R., & Marengoni, M. (2019, October). A survey of transfer learning for convolutional neural networks. In 2019 32nd SIBGRAPI conference on graphics, patterns, and images tutorials (SIBGRAPI-T) (pp. 47-57). IEEE.

[26] Butt, M. M., Iskandar, D. A., Abdelhamid, S. E., Latif, G., & Alghazo, R. (2022). Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features. *Diagnostics*, *12*(7), 1607.

[27] Mohammed, A. S., Hasanaath, A. A., Latif, G., & Bashar, A. (2023). Knee Osteoarthritis Detection and Severity Classification Using Residual Neural Networks on Preprocessed X-ray Images. *Diagnostics*, *13*(8), 1380.

[28] Latif, G., Abdelhamid, S. E., Mallouhy, R. E., Alghazo, J., & Kazimi, Z. A. (2022). Deep Learning Utilization in Agriculture: Detection of Rice Plant Diseases Using an Improved CNN Model. *Plants*, *11*(17), 2230.