

Analyzing Identification Friend or Foe (IFF) of Armored Vehicles: A Comparative Approach with Transfer Learning and Pre-Trained Models

Michael D. Parrish¹, Steven M. Corns², and James H. Schreiner¹

¹United States Military Academy, West Point, NY

²Missouri University of Science & Technology, Rolla, Mo

Corresponding Author's email: mdph2d@mst.edu

Author Note: COL(R) Michael Parrish is a former Professor of Practice at the United States Military Academy, West Point. Dr. Steven Corns is an Associate Professor of Engineering Management and Systems Engineering and Associate Chair of Graduate Studies at Missouri University of Science & Technology. COL James Schreiner is an Academy Professor in the United States Military Academy (USMA), and Associate Professor in the Department of Systems Engineering (DSE) where he currently serves as the Director of the Engineering Management Program.

Abstract: In military operations, the proper use of force in compliance with the laws of war is non-negotiable. As the battlespace is becoming more and more volatile, uncertain, complex, and ambiguous, military leaders are becoming more dependent upon advanced technologies such as artificial intelligence to assist in their decision-making process. One such area which military leaders need advanced technology support is positive control of all friendly assets as well as identification of enemy and noncombatant assets in all domains of air, land, sea, and space through the identification of friend or foe (IFF). The purpose of this study is to assess the viability of deep transfer learning to assist in military decision-making and answer the research question: Can pre-trained deep learning models be used to adequately identify and classify enemy targets? This study is designed as a comparative study to assess and compare various pre-trained deep-learning models to determine if they are adequate for targeting and engaging the enemy. An ensemble model was also incorporated using three pre-trained models and compared to the results of the individual models. A discussion of Human-in-the-loop concepts as well as the ethical considerations of the use of AI for IFF is incorporated in this study.

Keywords: Deep Transfer Learning, Pre-Trained Deep Learning Models, Target Classification, AI Ethics

1. Introduction

In military operations, the proper use of force in compliance with the laws of war is non-negotiable. Inherent in applying these laws, military leaders prioritize the proper employment of weapon systems against enemy combatants. From the strategic to the tactical level, military leaders have formal processes to ensure compliance with the proper use of force. A key component of these processes is rigorous attention to assessing potential targets before engagement.

The decision-making environment of the military leader is characterized as Volatile, Uncertain, Complex, and Ambiguous (VUCA) (Clark et al., 2010). In addition, time is limited for the commander to synthesize the plethora of data available and decide upon an appropriate course of action. It is critical that, in today's complex battlespace, military leaders have positive control of all friendly assets as well as identification of all enemy and noncombatant assets in all domains of air, land, sea, and space. This is accomplished using intelligent technologies such as IFF (Identify Friend or Foe) systems. The Department of Defense describes IFF systems as "sensitive identification devices that emit signals used to identify whether a platform is friendly or unknown to help prevent fratricide" (DSCA, 2018). This type of technology has become standard use via a transponder which is incorporated in all military as well as civilian aircraft for positive identification and control by air traffic controllers worldwide as well as ground and sea-based systems such as the Patriot Air Defense system, Aegis-equipped ships, and command and control platforms such as AWACS.

There are two types of target planning methodologies – deliberate and dynamic targeting. Deliberate targeting is characterized by preplanning and scheduled prosecution of given targets and is usually done at the operational and strategic levels. Dynamic targeting, on the other hand, applies to immediate targets of opportunity or surprises that are usually executed

at the tactical level such as air-air combat or direct ground engagement of the enemy. It is these dynamic targets of opportunity that the IFF systems are most critical to assist the individual pilot or commander in making immediate decisions about whether to engage a target or not.

Historically, the level of trust in technology has often influenced human errors. For instance, military leaders have sometimes disregarded information from technology due to distrust. This disregard has led to catastrophic failures in instances, and the emergence of AI further complicates this dynamic. One such example is the 1988 shooting down of Iranian Airlines Flight 655 by the USS Vincennes was a direct result of the failure of the Captain and crew of the Vincennes to properly use their Aegis AWS system resulting in their misidentification of the civilian aircraft. In that incident, it was found that had the crew utilized their Aegis in fully autonomous mode, the shootdown might not have ever occurred. Even in the semi-autonomous mode where the human makes the decision to fire, the Aegis had properly identified the target as civilian, but the Captain made the decision to fire manually thinking the airliner was an inbound hostile Iranian F-14, not trusting the information presented by the technology (Vincze, 2020).

Traditionally, at the tactical level, soldiers are expected to visually identify and confirm an enemy target before engaging their weapon systems. In previous wars, it was fairly easy to identify visually whether a combat vehicle was enemy or friendly because the systems were manufactured in their sphere of influence such as NATO or the Warsaw Pact where each had unique silhouettes. Such visual confirmation is becoming increasingly more difficult since combat vehicle types are exchanged across borders under foreign military or direct commercial sales agreements globally around the world (Gilman et al., 2014). This is further compounded by the addition of civilian vehicles used in military operations adding increased uncertainty as to who is the enemy. The use of artificial intelligence (AI) tools to help military leaders make tactical engagement decisions under this uncertainty is becoming more and more relevant to AI integration and accessibility in today's environment.

Image classification and object detection has become a popular AI task that is accomplished using deep learning algorithms and convolutional neural networks (CNN). Several popular CNNs have been provided to the public as open-source pre-trained models. The use of these pre-trained models for a given classification task using a different dataset of images than what the model was originally trained on is a machine learning technique called Transfer Learning. Such transfer learning enables the user to adopt a pre-trained model to build upon the existing knowledge to quickly complete the given task without having to develop and train a new model from scratch, saving time and improving performance.

However, as artificial intelligence continues to mature, human trust in technology will continue to be a major concern. A recent study found that the use of AI improved performance but did not change human risk-taking behaviors and that participants undervalued the AI recommendations in uncertain situations (i.e. lack of trust) (Elder et al., 2022). This trust concern is compounded in military decision-making when choosing to engage a potential enemy target using an AI-enabled weapon system. As will be discussed later in this study, the concept of Human-In-The-Loop (HITL) is a methodology to improve this trust relationship.

The purpose of this study is to assess the viability of deep transfer learning to assist in military decision-making and answer the research question: Can pre-trained deep learning models be used to adequately identify and classify enemy targets? To accomplish this, the study is designed as a comparative study to assess and compare various pre-trained deep-learning models to determine if they are adequate for targeting and engaging the enemy.

2. Related Research

As AI and machine learning become more mainstream in industry, there have been various research efforts focused on military use cases including IFF and targeting. Significant attention has been given to the use of AI technologies and techniques in both offensive and defensive operations (Layton, 2021). It is within the decision-making phase when AI technology has the greatest impact. In the early stages of development, AI systems have been used to forecast the enemy's actions by analogy with a historical database to create a predictive model. AI also assists in the process of decision-making under uncertainty. Expert systems have been developed to model human decision-making and transfer this knowledge to a machine where it can be applied consistently in each circumstance (Lesinski, et al., 2016; Schwartz, et al., 2020). These systems help to reduce the cognitive load of decision-makers by automating simple decisions, allowing a human operator to focus on more complex problems. The long-term goal is to have AI automate the entire decision-making process (Zhang, et al., 2020; Abaimov and Martellini, 2020; Billing, et al, 2021). Foreign adversaries, specifically China, have been researching and addressing the future of AI use in military warfare (Kania, 2022).

2.1 Use of AI for Military Targeting

Advances in machine learning in decision trees, clustering, and reinforcement learning have created increasingly intelligent systems and ultimately a force multiplier for the capabilities of the warfighter (Layton, 2021; Davis, 2022). Deep learning models have been used to increase the efficacy of automatic target recognition (ATR) systems (Kong, et al., 2022; Neupane and Seok, 2020; Janakiramaiah, et al., 2021). Convolutional neural networks (CNN) have been shown to be effective for ATR in SAR and moving target indication (MTI) sensors (Majumder et al., 2020; Feng, et al., 2021; Oveis, et al., 2022).

2.2 Deep Learning for Weapon Systems

Much of the literature describes the use of deep learning to improve Automatic Target Recognition (ATR) using Synthetic Aperture Radar (SAR) in the Radio Frequency (RF) spectrum (Majumder et al., 2020; Feng, et al., 2021). Several studies conducted involved training various CNN models as well as modifying popular (CNN) models to evaluate the efficacy of deep learning to improve ATR (Kong, et al., 2022; Neupane and Seok, 2020). Deep learning iterations continue to increase the level of ATR capability, an example being the research at UCSD in which a team used experimental methodology to show that a recurrent neural network could outperform a human in an online target detection task, the first to demonstrate the superiority of an algorithmic ATR system (Ma, et al., 2022). A more recent application of deep learning in ATR is a multi-task model that can identify, segment, and estimate the pose of the object (Jug et al.2022). Other multi-task studies included the use of Generative Adversarial Networks (GAN) for ATR (Zhuang, et al., 2024; Oh and Kim, 2021; Zhou et al., 2021).

2.3 AI IFF

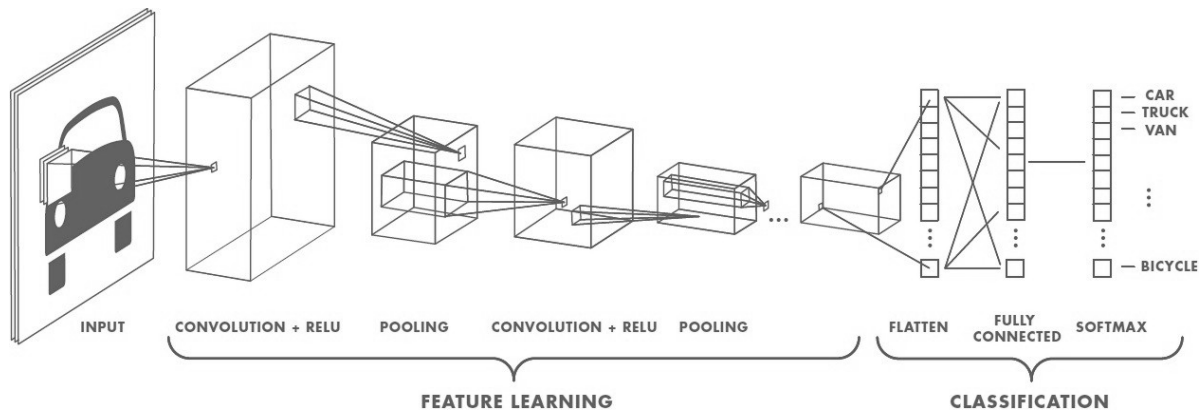
The significance of AI in combat identification has been emphasized and is growing, with studies focusing on the identification of friend or foe (IFF) systems using deep learning techniques (Ozmen & Aksoy, 2023). Additionally, the application of deep learning for military entity recognition and heterogeneous military entity recognition has been investigated, showcasing the fusion of deep learning and machine learning for enhanced recognition capabilities (Janakiramaiah, et al., 2021, Kong et al., 2022, Li et al., 2022, Liu et al., 2020). While multiple studies are using pre-trained models and transfer learning to improve the identification of various images, (Alhudhaif et al., 2022, Bhatt & Ganatra, 2022, Tamboli et al., 2023), the literature is relatively silent on its use to classify a military vehicle as a friendly asset or enemy target.

3. Methodology

To answer the research questions posed here, this study investigates Transfer Learning techniques to properly classify a dataset of military armored vehicles with various pre-trained models. Several of the most popular models for image classification were identified and used for this comparison study. The results focus on the validation accuracy of each model to correctly classify a vehicle image as either a friendly tank or an enemy (foe) tank. Finally, an ensemble model was developed and trained to compare its results to those of the individual models.

3.1 Pre-trained Deep Learning Image Classification Models Considered

It is well established that Convolutional Neural Networks (CNNs) have outperformed all previous machine learning methods for the task of object and image recognition. A CNN automatically learns a hierarchical representation of an image in a low-dimensional feature space, which allows for easy manipulation and understanding of the image. CNNs are designed to require minimal preprocessing and have shared weights, which drastically reduces the memory footprint and increases statistical efficiency. A typical CNN is composed of a combination of convolutional layers, non-linear activation functions, and a form of spatial pooling or normalization followed by several fully connected layers at the end. The major computational burden in training a CNN arises from the requirement to train the many parameters of the convolutional and fully connected layers from a large image dataset. Figure 1 shows graphically a typical CNN for image classification.



Source: [When to use Convolutional Neural Networks \(CNN\)? \(opengenius.org\)](https://opengenius.org)

Figure 1: Convolution Neural Network

This study used several popular pre-trained CNN models to train and assess their viability to accurately classify an image as a friendly or enemy (foe) armored vehicle. Each was pre-trained using the ImageNet dataset (TensorFlow, 2024). The details of the models considered are shown in Table 1.

Table 1 – Pretrained Model Details

Model	Size (MB)	Parameters (Millions)	Layers
MobileNetV2	9	2.3	154
ResNet50	90	23.6	175
VGG16	56	14.7	16
InceptionV3	83	22	311
DenseNet169	27	7	427
Xception	80	21	132
NASNetLarge	324	85	1039
EfficientNetV2B0	23	6	270

3.2 Data preprocessing

The datasets utilized for this study were images of various armored vehicle weapon systems from around the world and were obtained at www.Kaggle.com. These 1066 images were saved as .jpg files and sorted into two data folders. All Western-made equipment was incorporated into the “friend” data folder consisting of 497 files, while Warsaw Pact-based equipment was incorporated into the “foe” data folder consisting of 569 files. Sample images are shown in Figure 2.



Figure 2 – Example of Dataset Files by Class

Splitting the .jpg files into training, validation, and test datasets using 80/20 training/validation split resulted in a breakdown of files as follows:

```

Found 1066 files belonging to 2 classes.
Using 853 files for training.
Found 1066 files belonging to 2 classes.
Using 213 files for validation.
Number of validation batches: 6
Number of test batches: 1

```

3.3 Data Augmentation

Because the dataset is relatively small, the model incorporated data augmentation to generate additional images to increase the dataset size. This was accomplished by adding the Keras preprocessing layers or RandomFlip (set to 'horizontal'), RandomRotation (set to 0.2), and RandomZoom (set to 0.1) to randomize angular shifts in the existing files to generate new images resulting from the different angles of the original image. An example is shown in Figure 3.



Figure 3 – Example of Generated files due to Data Augmentation

3.4 Model Development

The base model was created by downloading the previously described most popular pre-trained image classification models. This base model is then customized for transfer learning following the established process described in TensorFlow and Keras. To accomplish this, the base layers of the pre-trained model are initially frozen and new feature extraction classification layers are added on top of the frozen base layers. These customized top layers were fixed for each pre-trained model use.

Normalization was accomplished as part of model development using the preprocessing input applications in Keras for each specific pre-trained model. Additionally, since overfitting and underfitting are common issues with transfer learning, the models incorporated early stopping and dropout layers to account for regularization.

The models were then compiled using an initial learning rate of 0.0001 with an initial training for 25 epochs. At this initial step, each model produces the baseline initial loss and accuracy rates as well as initial training loss and validation accuracy rates. The models are then further improved using fine-tuning techniques by unfreezing various top layers of the base model and retraining both new classification layers and several base model classification layers. The fine-tuning learning rate was reduced by a factor of 10 to 0.00001 and training was conducted for 25 epochs.

The fine-tuned loss and validation accuracy are determined followed by evaluating the model's performance using test images to establish prediction accuracy of the model to correctly identify if a given weapon system is a friend or foe vehicle. This process is repeated for each of the pre-trained models considered in this study.

3.5 Ensemble Deep Learning

To further evaluate various AI techniques to improve decision-making, an ensemble model was developed using three pre-trained models, VGG16, ResNet50, and MobileNetV2. While any combination and amount of the individual models could be chosen for inclusion in the ensemble model, these three individual models were selected because they are among the most popular pre-trained models used for image recognition and classification. The detailed descriptions of these models are detailed below.

3.5.1 VGG16

Developed by the Visual Graphics Group at Oxford University in 2014, this model consists of 16 convolutional layers with 3 fully connected layers as shown in Figure 4. It is a much larger model than most with 138 billion parameters and uses the ReLU activation function with a default image input size of 224 x 224. Because of its size, it is susceptible to overfitting but is one of the most popular pre-trained models for transfer learning for image classification (Simonyan and Zisserman, 2014).

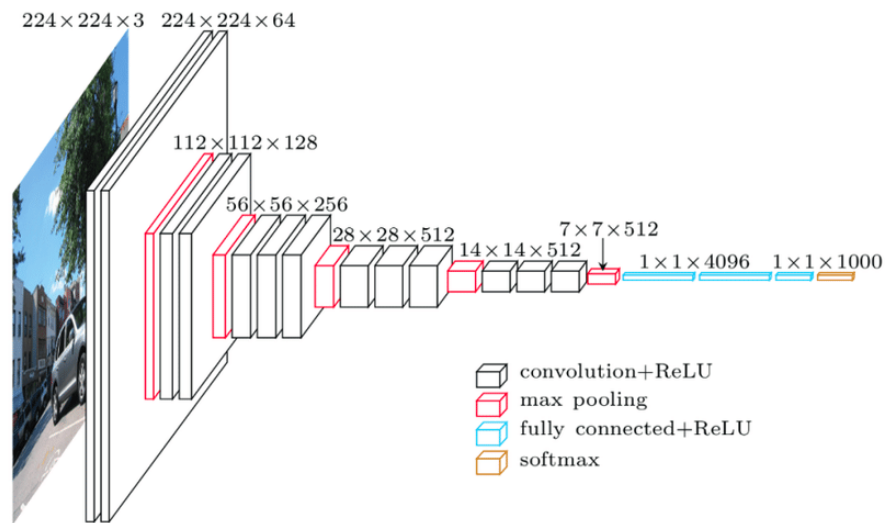


Figure 4 – Architecture of VGG-16 Model (Bezdan & Bačanin Džakula, 2019)

3.5.2 ResNet50

Residual Network (Fig. 5), developed in 2015 by Microsoft Corporation and known as ResNet50, is a deep convolutional neural network model consisting of 50 layers across five stages and uses the SoftMax activation function. Unlike the other models, it uses residual connections, also known as skip connections, enabling inputs to skip various layers thereby reducing training time. It was designed to improve accuracy by addressing the vanishing gradient issue present in other models such as VGG16 (Ali et. al., 2021).

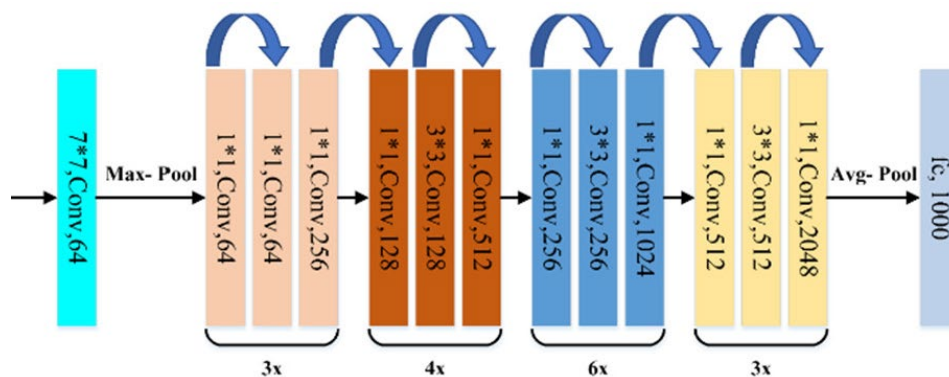


Figure 5 – Architecture of ResNet50 Model (Ali et al., 2021)

3.5.3 MobileNet V2

MobileNet (Fig. 6) was introduced in 2017 by Google as a streamlined CNN model intended for use with mobile applications. It is much smaller than other pre-trained models with its intent to improve efficiency by reducing computation time and size while maintaining comparable accuracy as compared to the deeper models (Howard et al., 2017).

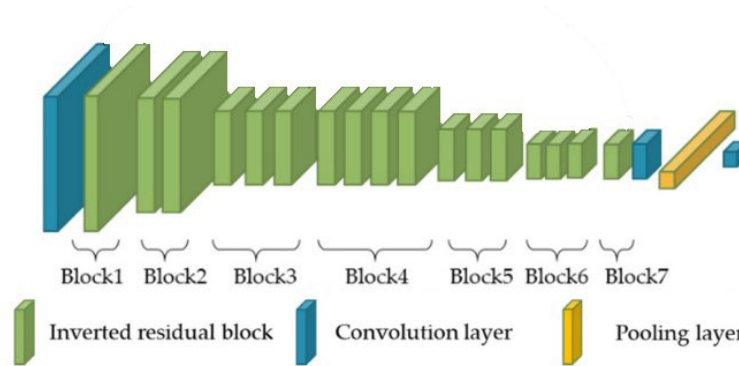


Figure 6 – Architecture of MobileNet V2 Model (Lin, et al., 2020)

The ensemble model averages each of the pre-trained models and its performance is compared to the performance of the individual models. Because the accuracy of each individual model was generally similar, the weights of each input were equal in the base ensemble model. The learning rate and initial epochs were the same for the ensemble model as they were for the individual model tests.

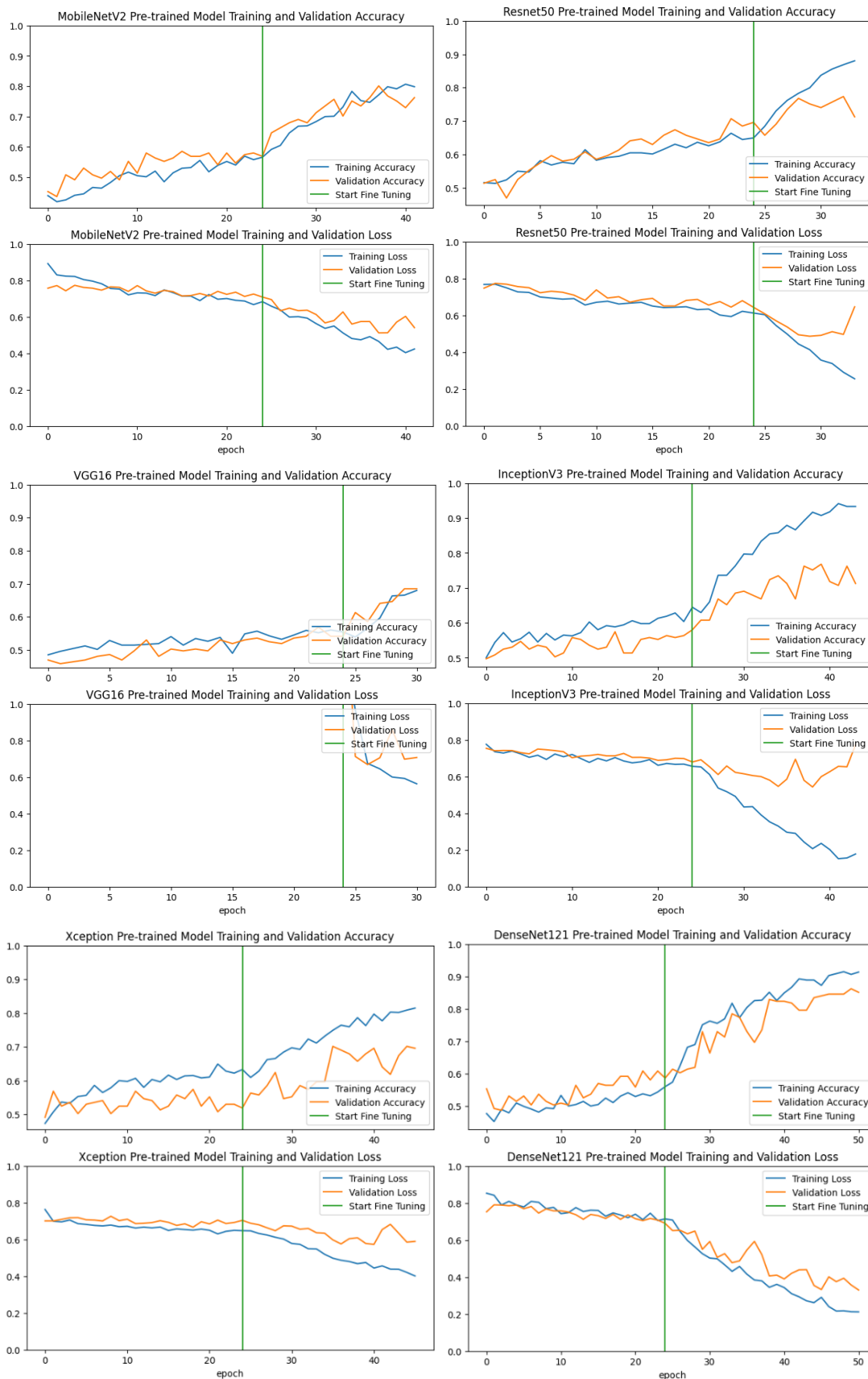
A final test was completed as a weighted average ensemble model where the weights of the input models were adjusted based on their accuracy by increasing the weight for the individual model with the highest accuracy and reducing the weight for the individual model that performed the worst of the three.

4. Results

The combined results of the individual pre-trained models across the phases of model training are provided in Table 2 and depicted graphically in Figure 7.

Table 2 – Individual Model Results

Model	Initial		Train 25 Epochs				Fine Tuning 25 Epochs				Test Accuracy	Early Stopping @ Epoch
	Loss	Accuracy	Loss	Accuracy	Val. Loss	Val. Accuracy	Loss	Accuracy	Val. Loss	Val. Accuracy		
MobileNetV2	0.8	0.51	0.7	0.61	0.67	0.61	0.35	0.84	0.49	0.76	0.78	Yes - 43
Resnet50	0.77	0.54	0.61	0.68	0.62	0.68	0.22	0.89	0.62	0.73	0.72	Yes - 34
VGG16	1.97	0.5	1.34	0.56	1.81	0.54	0.56	0.68	0.71	0.69	0.84	Yes - 30
InceptionV3	0.80	0.40	0.66	0.65	0.68	0.58	0.18	0.93	0.77	0.71	0.59	Yes - 43
DenseNet121	0.73	0.55	0.71	0.56	0.69	0.59	0.21	0.91	0.33	0.85	0.91	No
Xception	0.69	0.58	0.65	0.63	0.70	0.51	0.40	0.51	0.59	0.70	0.66	Yes - 45
NASNetLarge	0.75	0.43	0.62	0.63	0.67	0.60	0.07	0.58	0.69	0.55	0.81	Yes - 39
EfficientNetV2B0	0.70	0.50	0.64	0.64	0.65	0.61	0.25	0.90	0.52	0.75	0.75	Yes - 48



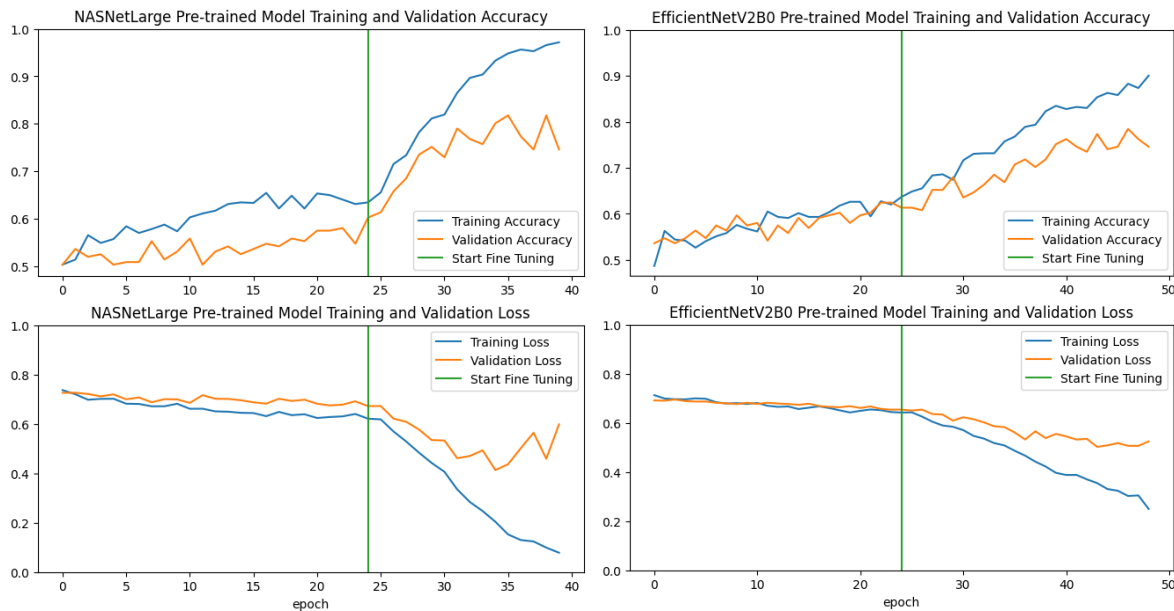


Figure 7 – Loss and Accuracy Graphs

The DenseNet121 pre-trained model proved to be the most accurate of all the individual models with a validation accuracy of 85% and test accuracy of 91%. This model was the only one to complete the full 50 epochs without early stopping due to overfitting. The validation and test accuracy for all the models is in line with previous studies using pre-trained models for transfer learning.

The ensemble accuracy showed a significant improvement of validation accuracy of 12% on average over the individual results of three pre-trained models used (.76, .73, .69) as shown in Table 3.

Table 3 – Ensemble Model Results

Model	Loss	Accuracy	Val. Loss	Val. Accuracy	Test Accuracy
Ensemble	0.02	0.99	0.57	0.81	0.72

Finally, adjusting the individual weights based upon the validation accuracy of the individual model inputs from equal weighting to $w_1 = .35$, $w_2 = 0.33$, and $w_3 = .32$, yielded a further 6% validation accuracy improvement over the base ensemble model as shown in Table 4 below.

Table 4 – Weighted Average Ensemble Model Results

Model	Loss	Accuracy	Val. Loss	Val. Accuracy	Test Accuracy
Weighted Ensemble	0.09	0.97	0.62	0.86	0.81

5. Discussion

5.1 Limitations and Constraints

While transfer learning is an expedient way to quickly train a model using established pre-trained models to generalize results, this approach has several limitations that affect the optimum performance and outcomes of the desired task being performed. In particular, the choice of which pre-trained model to use may not be well-suited to the data being evaluated which, in itself, could be an insufficient amount for the pre-trained model to be effective. This limitation applies to this study in that the number of tank images used was 1066 where data augmentation was necessary as described in the preprocessing discussion above. Additionally, it is common with transfer learning to overfit the model during the fine-tuning phase. This was observed in this study of seven of the eight pre-trained models as evidenced by the early stopping as shown in Table 2 above. Finally, this study was limited to evaluating eight of the most common pre-trained models available in Tensorflow.

Although this study demonstrated the potential for utilizing pre-trained models to distinguish between friendly and enemy tanks, their accuracy seems insufficient for many military combat situations. US military rules of engagement require military leaders to identify weapons control status for any operation defined as follows:

- WEAPONS HOLD. Engage only if engaged or ordered to engage.
- WEAPONS TIGHT. Engage only targets positively identified as enemy.
- WEAPONS FREE. Engage targets not positively identified as friendly. (Army MCoE, 2024)

In most cases on the battlefield, the rules of engagement require positive identification of an enemy, so anything less than 95% accuracy would be insufficient, and, in some cases (such as an urban environment), 100% accuracy would be expected. However, in a weapons-free scenario, where combatants anticipate all encountered vehicles will be enemies, the classification of friend versus foe becomes unnecessary. One way to improve this model is to incorporate Human-in-the-loop deep learning.

5.2 Human-in-the loop

Human-in-the-loop (HITL) deep learning is a well-researched approach to machine learning that incorporates human expertise to assist in the training and validation of the Machine Learning/Deep Learning (ML/DL) models. This human-machine interface is the key to the goal of improving the accuracy and efficiency of both machine-learning models and humans. There are three general approaches depending upon the level of control of the learning process as follows:

- **Active Learning (AL)** – AL is a semi-supervised learning process where the model iteratively uses both labeled and unlabeled training data. In general, a sampling strategy is selected whereby the model iteratively trains on a given dataset to identify the unlabeled data point for the human expert to properly label. This process continues until the desired end state is achieved. The model controls the learning process while the human serves as the expert “oracle” to properly annotate data. AL is best used in situations where labeling of data is expensive or time-consuming such as image classification and natural language processing (NLP).
- **Interactive Machine Learning (IML)** – IML builds upon the concept of AL but, in this case, control of the learning process is shared between the model and the human as both have tasks that they are more efficient at performing to accomplish actively during the training which is different than the passive interface of classical machine learning. Key to this approach is the human-machine interface (think human-centered design).
- **Machine Teaching (MT)** – For MT, the human serves as the teacher to teach the model to an algorithm with the goal of transferring human knowledge to the machine. In this approach, the human subject matter expert does not necessarily need to have knowledge of machine learning. However, a requirement and a risk of this approach is that the human teacher decides on the example sets as well as the curriculum (level of difficulty) for the training.

One of the many strengths of HITL deep learning is its ability to take raw, unstructured data and automatically discover the relationships necessary to detect or classify images, NLP, etc. As such, HITL-DL is a viable approach to improving AI models (Mosqueira-Rey, et al., 2023).

Although deep learning is becoming more prevalent as a very effective AI method, deep learning continues to struggle with large and complex data set issues such as bias and modeling errors resulting in overfitting (training error is small but validation/test error is large), and noise in the data which could cause deep learning hallucinations where the resultant output is nonsensical. Adding HITL-DL helps resolve those concerns. Figure 8 shows the steps in machine learning for HITL knowledge to be inserted for algorithm optimization (Wu, X. et al., 2022).

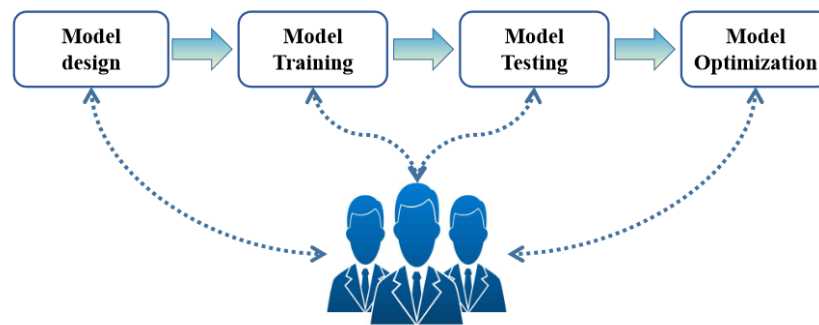


Figure 8 – Human-in-the-Loop System Machine Learning (Wu, X. et al., 2022)

Another area where deep learning is currently used, yet continues to struggle, to replicate real-world scenarios is autonomous driving. A recent study showed how real-time HITL-DL, called human guidance-based deep reinforcement learning (Hug-DRL), improved learning efficiency and training performance as compared to other reinforcement learning techniques (Wu, J. et al, 2023). Other areas, such as the medical field where patient safety is paramount, are incorporating HITL as a critical component to improving human trust in AI. For these ongoing studies, such as in human movement, medicine, and warfare, the debate on the ethical use of AI to directly impact human life has just begun.

5.3 Ethical Considerations

Beyond the need for the Human-in-the-Loop to use pre-trained models to improve the proper classification of military vehicles, it is important to understand the ethical considerations of AI and autonomous weapon systems on the battlefield.

Ethics are “the moral principles that govern a person’s behavior” (Oxford Dictionary, 2023). When reviewing such principles, the two most common philosophical approaches to establishing ethical limitations are the study of the means and the study of the ends of an action. More specifically, in Kantian ethics, also known as Deontological ethics, the action taken is morally more important than the result of the action. This belief is the basis for our established view that lying is morally wrong regardless of the benefits achieved by such a lie (Kant, 1785). Teleological ethics, also known as utilitarian ethics, on the other hand, focuses on the outcome of an action to establish the moral value of such action taken (Ismail and Abderrahmane, 2018). These two theories form the basis for determining the ethical limitations of the models and approaches being implemented.

When discussing the ethical limitations of IFF related to military operations, we must first understand the independent and internationally accepted legal principles of Jus ad Bellum, the justification for a state to enter into war, and Jus in Bello, the rules governing the conduct of hostilities. Every military officer is educated in the ethics of warfare. The established humanitarian protocols of the Geneva Convention are a part of Jus in Bello. The Department of Defense’s Law of War manual discusses the concept of proportionality as follows:

In jus ad bellum, proportionality refers to the principle that the overall goal of the State in resorting to war should not be outweighed by the harm that the war is expected to produce. However, the principle of proportionality in jus in bello generally refers to the obligations to take feasible precautions in planning and

conducting attacks and to refrain from attacks in which the expected loss of civilian life, injury to civilians, and damage to civilian objects incidental to the attack would be excessive. Therefore, although a jus ad bellum proportionality analysis might consider the harm suffered by enemy military forces in the fighting, a jus in bello proportionality analysis would not (DOD, 2023, p. 87).

Furthermore, DOD's manual discusses the viability and legality or prohibition of new weapons systems utilizing a three-part test as follows:

- whether the weapon's intended use is calculated to cause superfluous injury;
- whether the weapon is inherently indiscriminate; and
- whether the weapon falls within a class of weapons that has been specifically prohibited. (DOD, 2023, p. 347)

The manual, paragraph 6.4.2, further defines currently prohibited weapons systems due to active treaties the United States is party to, or current international laws as follows:

- poison, poisoned weapons, poisonous gases, and other chemical weapons;
- biological weapons;
- certain environmental modification techniques;
- weapons that injure by fragments that are non-detectable by X-rays;
- certain types of mines, booby-traps, and other devices; and
- blinding lasers. (DOD, 2023, p. 351)

The manual also discusses the applicability of autonomous systems and even highlights that target selection is an authorized capability as described in paragraph 6.5.9.1:

Other weapons may have more sophisticated autonomous functions and may be designed such that the weapon is able to select targets or to engage targets automatically after being activated by the user. For example, the United States has used weapon systems for local defense with autonomous capabilities designed to counter time-critical or saturation attacks. These weapon systems have included the Aegis ship defense system and the Counter-Rocket, Artillery, and Mortar (C-RAM) system (DOD, 2023, p. 362).

It should be noted that the weapon systems listed above include IFF capabilities to identify targets before autonomously engaging a hostile threat.

Current ethical issues being discussed by the international community include topics such as if the use of AWS would make war more likely and circumvent Jus ad Bellum or would terrorist organizations using AWS unjustly violate Jus in Bello. Perhaps the most concerning ethical question is the autonomy versus the autonomous nature of AWS in that future cognitive capabilities of such systems could enable them to autonomously decide when to start a war (Taddeo and Blanchard, 2023). Such a hypothetical and unlikely scenario has not been taken seriously, yet it must be considered for those nation-states that do not align with Western moral codes and laws.

Similarly, as many military AI systems are dual-use, ethical limitations of the use of AI apply to this study. Emerging considerations are related to the proper use of AI technologies. At the national level, the US government is leading several initiatives on determining the proper governance and oversight of AI with participation from federal agencies, academia, and industry. The US Department of Defense has established five guiding principles in dealing with ethics as follows:

1. **Responsible.** DoD personnel will exercise appropriate levels of judgment and care while remaining responsible for the development, deployment, and use of AI capabilities.
2. **Equitable.** The Department will take deliberate steps to minimize unintended bias in AI capabilities.
3. **Traceable.** The Department's AI capabilities will be developed and deployed such that relevant personnel possess an appropriate understanding of the technology, development processes, and operational methods applicable to AI capabilities, including transparent and auditable methodologies, data sources, and design procedure and documentation.

4. **Reliable.** The Department's AI capabilities will have explicit, well-defined uses, and the safety, security, and effectiveness of such capabilities will be subject to testing and assurance within those defined uses across their entire life cycles.
5. **Governable.** The Department will design and engineer AI capabilities to fulfill their intended functions while possessing the ability to detect and avoid unintended consequences, and the ability to disengage or deactivate deployed systems that demonstrate unintended behavior. (Department of Defense, 2020, February 24)

The standardization of ethical considerations in the use of AI is a relatively new area of research and includes ongoing philosophical discussions related to the ethical considerations of developing sentient AI systems. If such systems could become self-aware, can they be considered a person? And if so, do they get legal status and rights? Are they to be treated as colleagues? Fortunately, the majority consensus is that such capabilities are far in the future, but the concern is relevant today, especially given the earlier discussions related to AWS (Löhr, 2023).

While American law has established the legality of using Autonomous Weapon Systems (AWS), there remains an ongoing international debate regarding the ethical use of such weapons in a fully autonomous capacity. As there is no agreed-upon approach, the current thought is that there must be Meaningful Human Control (MHC) over all weapon systems where humans exercise moral control, especially those used offensively (Amoroso and Tamburrini, 2020; Diggelen et al., 2023).

6. Conclusion

Throughout the history of war, technological advances have resulted in changes to how military operations are conducted. Currently, there is significant work being done related to technology that will impact the future of the offensive and defensive nature of warfare based upon current events such as the effective use of drones in Ukraine as well as drone swarm tactics as was recently attempted in Israel. In fact, Ukraine is already using semi-autonomous AI drones to detect (IFF) and engage enemy targets (Euromaidan Press, 2023, October 6).

This study demonstrated that utilizing a weighted-average ensemble model provides improved results as compared to individual pre-trained models for military tank classification. However, for military decision-makers, this study showed that, while transfer learning with pre-trained models can distinguish between friendly and enemy tanks, their accuracy seems insufficient for many military combat situations. Further studies are required to develop and evaluate models to achieve appropriate human-AI trust before endorsing this methodology for military decision-makers to use as an effective IFF tool. Such research should include developing a Human-In-The-Loop approach to address the ethical concerns of autonomous weapon systems. Additional studies can incorporate larger datasets with additional armored vehicles as well as evaluation of additional pre-trained as well as self-developed models.

We are living in a time where AI as well as other technology and process advances will soon become an inflection point of significant social and military change as has been done throughout human history. Leaders must embrace this change and soon adjust how they make decisions.

7. References

- Abaimov, S., & Martellini, M. (2020). Artificial Intelligence in Autonomous Weapon Systems. In M. Martellini & R. Trapp (Eds.), *21st Century Prometheus: Managing CBRN Safety and Security Affected by Cutting-Edge Technologies* (pp. 141–177). Springer International Publishing. https://doi.org/10.1007/978-3-030-28285-1_8
- Alhudhaif, A., Saeed, A., Imran, T., Kamran, M., S. Alghamdi, A., O. Aseeri, A., & Alsubai, S. (2022). A Particle Swarm Optimization Based Deep Learning Model for Vehicle Classification. *Computer Systems Science and Engineering*, 40(1), 223–235. <https://doi.org/10.32604/csse.2022.018430>
- Ali, L., Alnajjar, F., Jassmi, H. A., Gocho, M., Khan, W., & Serhani, M. A. (2021). Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. *Sensors*, 21(5), 1688. <https://doi.org/10.3390/s21051688>
- Amoroso, D. and Tamburrini, G. (2020). Autonomous Weapons Systems and Meaningful Human Control: Ethical and Legal Issues. *Current Robotics Reports* 1, no. 4 (December 1, 2020): 187–94. <https://doi.org/10.1007/s43154-020-00024-3>.

- Bhatt, A., & Ganatra, A. (2022). Explosive Weapons and Arms Detection with Singular Classification (WARDIC) on Novel Weapon Dataset using Deep Learning: Enhanced OODA Loop. *Engineered Science, Volume 20 (December 2022)*, 252–266. <https://doi.org/10.30919/es8e718>
- Bezdan, T., & Bačanić Džakula, N. (2019). Convolutional Neural Network Layers and Architectures. *Proceedings of the International Scientific Conference - Sinteza 2019*, 445–451. <https://doi.org/10.15308/Sinteza-2019-445-451>
- Billing, D. C., Fordy, G. R., Friedl, K. E., & Hasselstrøm, H. (2021). The implications of emerging technology on military human performance research priorities. *Journal of Science and Medicine in Sport, 24(10)*, 947–953. <https://doi.org/10.1016/j.jsams.2020.10.007>
- Clark, M., Allen, C., Keegan, T., Meinhart, R., Wong, L., & Reed, G. (2010). *Strategic leadership primer* (p. 0071). S. J. Gerras (Ed.). Department of Command, Leadership and Management, US Army War College. <https://doi.org/10.1037/e660342010-001>
- Davis, S. I. (2022). Artificial intelligence at the operational level of war. *Defense & Security Analysis, 38(1)*, 74–90. <https://doi.org/10.1080/14751798.2022.2031692>
- Defense Security Cooperation Agency. (2018). “Chapter 3 | Defense Security Cooperation Agency.” Accessed October 14, 2023. <https://samm.dsca.mil/chapter/chapter-3>.
- Diggelen, J., van den Bosch, K., Neerinx, M., and Steen, M. Designing for Meaningful Human Control in Military Human-Machine Teams. arXiv, May 12, 2023. <https://doi.org/10.48550/arXiv.2305.11892>.
- Elder, H., Shank, D., Canfield, C., Rieger, T., & Hines, C. (2022). Knowing When to Pass: The Effect of AI Reliability in Risky Decision Contexts. *Human factors*, <https://doi.org/10.1177/00187208221100691>
- “Ethic Noun – Definition, Pictures, Pronunciation and Usage Notes | Oxford Advanced American Dictionary at OxfordLearnersDictionaries.Com.” Accessed October 11, 2023. https://www.oxfordlearnersdictionaries.com/us/definition/american_english/ethic.
- Euromaidan Press. “Ukraine Uses Drones Equipped with AI to Destroy Military Targets,” October 6, 2023. <https://euromaidanpress.com/2023/10/06/ukraine-uses-drones-equipped-with-ai-to-destroy-military-targets/>.
- Feng, S., Ji, K., Zhang, L., Ma, X., & Kuang, G. (2021). SAR Target Classification Based on Integration of ASC Parts Model and Deep Learning Algorithm. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*, 10213–10225. <https://doi.org/10.1109/JSTARS.2021.3116979>
- Gilman, D., Nichols, R., & Totman, P. J. C. (2014). Foreign military sales. *Defense Security Cooperation Agency*. http://www.dsca.mil/sites/default/files/final-fms-dcs_30_sep.pdf.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications* (arXiv:1704.04861). arXiv. <http://arxiv.org/abs/1704.04861>
- Ismail, N. S. A., & Benlahcene, A. (2018). A Narrative Review Of Ethics Theories: Teleological & Deontological Ethics. *IOSR Journal Of Humanities And Social Science (IOSR-JHSS)*, 23, 31–38. <https://doi.org/10.9790/0837-2307063138>
- Janakiramaiah, B., Kalyani, G., Karuna, A., Prasad, L. V. N., & Krishna, M. (2021). Military object detection in defense using multi-level capsule networks. *Soft Computing, 27(2)*, 1045–1059. <https://doi.org/10.1007/s00500-021-05912-0>
- Jug, J., Lampe, A., Štruc, V., & Peer, P. (2022). *Body Segmentation Using Multi-task Learning* (arXiv:2212.06550). arXiv. <http://arxiv.org/abs/2212.06550>
- Kania, E., (2019). Chinese Military Innovation in the AI Revolution. *The RUSI Journal*, 164(5-6), 26-34. <https://doi.org/10.1080/03071847.2019.1693803>
- Kant, I. *Groundwork for the Metaphysics of Morals*. Edited by Thomas E. Hill and Arnulf Zweig. New York: Oxford University Press, 1785.
- Kong, L., Wang, J., & Zhao, P. (2022). YOLO-G: A Lightweight Network Model for Improving the Performance of Military Targets Detection. *IEEE Access, 10*, 55546–55564. <https://doi.org/10.1109/ACCESS.2022.3177628>
- Layton, P. (2021). *Fighting Artificial Intelligence Battles Operational Concepts for Future AI-Enabled Wars*.
- Lesinski, G., Corns, S. M., & Dagli, C. H. (2016). A fuzzy genetic algorithm approach to generate and assess meta-architectures for non-line of site fires battlefield capability. *2016 IEEE Congress on Evolutionary Computation (CEC)*, 2395–2401. <https://doi.org/10.1109/CEC.2016.7744085>
- Li, H., Yu, L., Zhang, J., & Lyu, M. (2022). Fusion Deep Learning and Machine Learning for Heterogeneous Military Entity Recognition. *Wireless Communications and Mobile Computing, 2022*, e1103022. <https://doi.org/10.1155/2022/1103022>
- Lin, Z., Ye, H., Zhan, B., & Huang, X. (2020). An Efficient Network for Surface Defect Detection. *Applied Sciences, 10(17)*, 6085. <https://doi.org/10.3390/app10176085>

- Liu, Y., Yu, Y., Wang, L., Nyima, T., Zhaxi, N., Huang, H., & Deng, Q. (2020). Classification of Tank Images Using Convolutional Neural Network. *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 210–214. <https://doi.org/10.1109/ITNEC48623.2020.9085151>
- Löhr, G. “If Conceptual Engineering Is a New Method in the Ethics of AI, What Method Is It Exactly?” *AI and Ethics*, May 16, 2023. <https://doi.org/10.1007/s43681-023-00295-4>.
- Ma, Q., Li, S., & Cottrell, G. W. (2022). Adversarial Joint-Learning Recurrent Neural Network for Incomplete Time Series Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(4), 1765–1776. <https://doi.org/10.1109/TPAMI.2020.3027975>
- Majumder, U. K., Blasch, E. P., & Garren, D. A. (2020). *Deep Learning for Radar and Communications Automatic Target Recognition*. Artech House.
- Mission Command. (n.d.). Retrieved April 27, 2024, from https://www.moore.army.mil/Infantry/DoctrineSupplement/ATP3-21.8/appendix_b/DirectFireControl/FireControlMeasures/index.html
- Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2023). Human-in-the-loop machine learning: A state of the art. *Artificial Intelligence Review*, 56(4), 3005–3054. <https://doi.org/10.1007/s10462-022-10246-w>
- Neupane, D., & Seok, J. (2020). A Review on Deep Learning-Based Approaches for Automatic Sonar Target Recognition. *Electronics*, 9(11), Article 11. <https://doi.org/10.3390/electronics9111972>
- Oh, J., & Kim, M. (2021). PeaceGAN: A GAN-Based Multi-Task Learning Method for SAR Target Image Generation with a Pose Estimator and an Auxiliary Classifier. *Remote Sensing*, 13(19), Article 19. <https://doi.org/10.3390/rs13193939>
- Oveis, A. H., Giusti, E., Ghio, S., & Martorella, M. (2022). Moving and Stationary Targets Separation in SAR Signal Domain Using Parallel Convolutional Autoencoders with RPCA Loss. *2022 IEEE Radar Conference (RadarConf22)*, 1–6. <https://doi.org/10.1109/RadarConf2248738.2022.9764168>
- Özmen, M., Aksoy, B. (2023). An Example Application for An Identification of Friend and Foe (IFF) System Appropriate for Unmanned Aerial Vehicles (UAV) Based on Deep Learning. *J Intell Robot Syst* **107**, 36 (2023). <https://doi.org/10.1007/s10846-023-01840-3>
- Schwartz, P. J., O’Neill, D. V., Bentz, M. E., Brown, A., Doyle, B. S., Liepa, O. C., Lawrence, R., & Hull, R. D. (2020). AI-enabled wargaming in the military decision-making process. *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*, 11413, 118–134. <https://doi.org/10.1117/12.2560494>
- Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv:1409.1556). arXiv. <http://arxiv.org/abs/1409.1556>
- Taddeo, M., & Blanchard, A. (2023). A Comparative Analysis of the Definitions of Autonomous Weapons. In F. Mazzi (Ed.), *The 2022 Yearbook of the Digital Governance Research Group* (pp. 57–79). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-28678-0_6
- Tamboli, S., Jagadale, K., Mandavkar, S., Katkade, N., & Ruprah, T. S. (2023). A Comparative Analysis of Weapons Detection Using Various Deep Learning Techniques. *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, 1141–1147. <https://doi.org/10.1109/ICOEI56765.2023.10125710>
- Transfer learning and fine-tuning | TensorFlow Core*. (2024). TensorFlow. Retrieved April 27, 2024, from https://www.tensorflow.org/tutorials/images/transfer_learning
- United States Army Maneuver Center of Excellence (MCoE). (2024). *Infantry Platoon and Squad (ATP 3-21.8)*. Retrieved from https://armypubs.army.mil/epubs/DR_pubs/DR_a/ARN40007-ATP_3-21.8-000-WEB-1.pdf
- U.S. Department of Defense. (2023). *Department of Defense Law of War Manual*.
- U.S. Department of Defense. (2020, February 24). *Ethical principles for artificial intelligence – ai.mil*. Ethical Principles for Artificial Intelligence. https://www.ai.mil/docs/Ethical_Principles_for_Artificial_Intelligence.pdf
- Vincze, V. (2020). The USS Vincennes incident: A case study involving Autonomous Weapon Systems. *Honvédségi Szemle – Hungarian Defence Review*, 148(2), Article 2. 92-101. <https://doi.org/10.35926/HDR.2020.2.6>
- When to use Convolutional Neural Networks (CNN)*. OpenGenusIQ. Retrieved March 26, 2024, from <https://iq.opengenus.org/when-to-use-convolutional-neural-network-cnn/>
- Wu, J., Huang, Z., Hu, Z., & Lv, C. (2023). Toward Human-in-the-Loop AI: Enhancing Deep Reinforcement Learning via Real-Time Human Guidance for Autonomous Driving. *Engineering*, 21, 75–91. <https://doi.org/10.1016/j.eng.2022.05.017>
- Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems*, 135, 364–381. <https://doi.org/10.1016/j.future.2022.05.014>
- Zhang, Y., Dai, Z., Zhang, L., Wang, Z., Chen, L., & Zhou, Y. (2020). Application of Artificial Intelligence in Military: From Projects View. *2020 6th International Conference on Big Data and Information Analytics (BigDIA)*, 113–116. <https://doi.org/10.1109/BigDIA51454.2020.00026>

Zhao, B., Wang, C., Fu, Q., & Han, Z. (2021). A Novel Pattern for Infrared Small Target Detection With Generative Adversarial Network. *IEEE Transactions on Geoscience and Remote Sensing*, 59(5), 4481–4492.

<https://doi.org/10.1109/TGRS.2020.3012981>

Zhuang, X., Li, D., Wang, Y., & Li, K. (2024). Military target detection method based on EfficientDet and Generative Adversarial Network. *Engineering Applications of Artificial Intelligence*, 132, 107896.

<https://doi.org/10.1016/j.engappai.2024.107896>