Developing a YOLOv8-Based Model for Automated Detection of Military Vehicles

Viktoriia Yuzkiv, Angelo Di Gianvito

Barcelona School of Economics, 2024

Motivation

In the digital age, the accurate and timely identification of military vehicles in news imagery is critical for a variety of applications, including media analysis, conflict monitoring, and historical archiving. By developing a robust model based on YOLOv8 for detecting military vehicles, we can significantly enhance the capabilities of automated systems to analyze media coverage of conflicts. This technology can be used to quantify the presence of military vehicles in news images, thereby serving as a proxy for understanding the intensity and focus of media reporting on wars and conflicts.

Literature Review

Extensive literature has covered applications of deep learning on war-related images. Weapon detection is one of the most prominent applications in the currently available literature. This is mainly applied to security applications like in “Weapon detection using artificial intelligence and deep learning for security applications”, or more specifically to CCTV image data for recognition like in the study of 2021 “Weapon detection in real-time cctv videos using deep learning”.

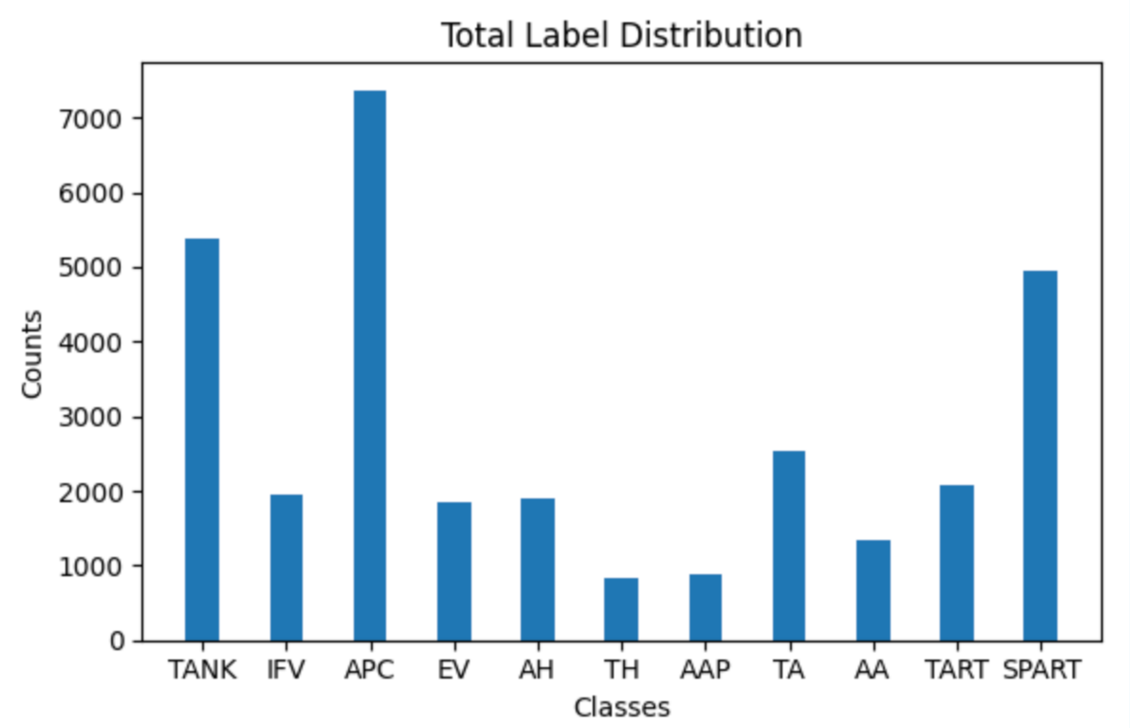
In this context, object detection models usually are the most commonly used technique to detect weapons in images for security purposes. The study "Battlefield image situational awareness application based on deep learning" used deep learning techniques to develop a situational awareness system for processing battlefield images. This was achieved utilizing a YOLO (You Only Look Once) object detection model to output confidence probabilities that objects in the images were helicopters, missiles, tanks, soldiers or guns.

Finally, in the paper “Machine Learning Based Analysis of Finnish World War II Photographers”, the authors use YOLO for object detection in historical photographs from the Finnish Civil War. It analyzes how photographers framed and captured images during wartime including tanks and other weapons.

Data

The dataset used in this project originates from the paper "Military Decision-Making Process Enhanced by Image Detection" (2023) by Nikola Žigulić, Matko Glučina, Ivan Lorencin, and Dario Matika. This dataset comprises 11,800 labeled images specifically designed for the YOLO detection algorithm. The images are categorized into the following classes: Tank (TANK), Infantry Fighting Vehicle (IFV), Armored Personnel Carrier (APC), Engineering Vehicle (EV), Assault Helicopter (AH), Transport Helicopter (TH), Assault Airplane (AAP), Transport Airplane (TA), Anti-Aircraft Vehicle (AA), Towed Artillery (TART), and Self-Propelled Artillery (SPART).

The dataset is divided into training, validation, and test sets to ensure comprehensive model evaluation and robust performance analysis. Looking at the classes distribution, we can see that APC and TANK are the most frequently occurring classes across all sets, while TH and AAP are the least represented.



Methodology

We utilized the YOLOv8 (You Only Look Once) object detection framework, known for its speed and accuracy in real-time object detection tasks. YOLO divides images into grids and predicts bounding boxes and class probabilities directly from full images in a single evaluation, making it highly efficient.

To fine-tune the YOLOv8 model on our dataset for better performance, we used the following hyperparameters: image size (640), number of epochs (50), and batch size (16). This fine-tuning process involved continuous training on the provided dataset, enhancing the model's ability to detect and classify military vehicles accurately.

Results

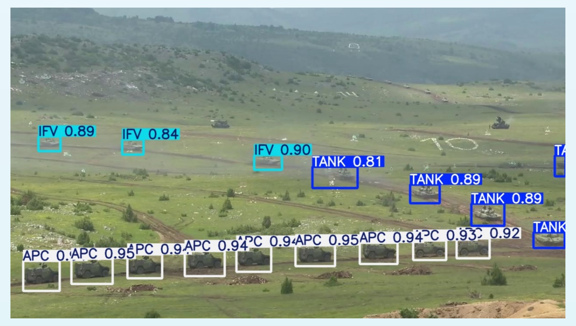
The fine-tuned YOLOv8 model demonstrated good performance. Training losses decrease steadily across epochs, indicating that the model is learning and improving over time. Both val/box_loss and val/cls_loss decrease over time, reflecting improvements in the model’s performance on the validation set as well.

On the test set our fine-tuned YOLOv8 model showed a strong performance across various classes with an overall precision of 0.971 and recall of 0.961, with mAP50-95 of 0.824. The highest performance was achieved on the class EV: 0.997 precision, 0.998 recall, 0.925 mAP50-95. Classes TANK and IFV also have high performance metrics. However, some classes such as AH and AAP exhibit lower performance, possibly due to fewer instances in the dataset. Moreover, the model performs slightly worse in scenarios where the vehicle is extremely small, obscured by smoke, or when the image resolution is very low. In such cases, the model's confidence may decrease, resulting in potential missed detections. However, these are instances where even human eyes find it challenging to detect the vehicle. Overall, the model demonstrates its ability to handle varied conditions and detect instances with high accuracy across different IoU thresholds, as reflected in the mAP50-95 metric.

Detailed performance on the test set across all classes can be seen on the table:

Class	Images	Instances	Box(P	R	mAP50	m
all	3234	7863	0.971	0.961	0.98	0.824
TANK	559	1213	0.981	0.988	0.994	0.843
IFV	217	426	0.986	0.986	0.995	0.857
APC	406	1686	0.959	0.98	0.986	0.853
EV	497	556	0.997	0.998	0.995	0.925
AH	169	496	0.931	0.816	0.933	0.642
TH	174	243	0.969	0.979	0.989	0.788
AAP	168	269	0.955	0.949	0.963	0.714
TA	310	664	0.997	0.998	0.995	0.922
AA	115	322	0.993	0.997	0.995	0.921
TART	436	596	0.966	0.955	0.979	0.859
SPART	568	1392	0.949	0.919	0.96	0.738

Additionally, the examples of predictions on the test set are presented below.



Applications

1. Military Decision Making

Military vehicles object detection technologies significantly impact decision-making processes within military operations. These technologies provide critical real-time intelligence and situational awareness, enabling commanders to make informed and timely decisions on the battlefield. By accurately identifying and tracking military vehicles such as tanks, infantry fighting vehicles (IFVs), and helicopters, these detection systems facilitate tactical advantage, force protection, target prioritization, mission planning.

2. Analysis of Conflict Coverage in Media/News

Military vehicles object detection technologies not only impact military operations but also allow for analysis of how conflicts and wars are covered in media and news outlets. These technologies provide valuable insights that contribute to comparative studies of conflict coverage and visual documentation of military activities.

Future Work

1. *Augmenting Data for Lower Represented Classes*

Augmenting data can enhance model robustness by expanding dataset diversity through additional image collection, data augmentation techniques, and synthetic data generation.

2. Analyzing Current Wars Coverage in News Media

Using the model to analyze media coverage of current wars offers opportunities to investigate temporal shifts in reporting styles, media bias, and the visual portrayal of military engagements.

3. Further Hyperparameter optimization

Hyperparameter optimization, including tuning learning rates, batch sizes, etc., could be performed for improving detection accuracy and efficiency.

References

Harsh Jain, et al. «Weapon detection using artificial intelligence and deep learning for security applications.»

Muhammad Tahir Bhatti, et al.. «Weapon detection in real-time cctv videos using deep learning.»

Peng et al. "Battlefield image situational awareness application based on deep learning»

Chumachenko et al. «Machine Learning Based Analysis of Finnish World War II Photographers»