Tiako P.F. (ed) *Intelligent and Sustainable Solutions*. Chronicle of Computing. OkIP. CAIF25#10
 https://doi.org/ 10.55432/978-1-6692-0011-6_7

Close-Combat Weapon Detection in Crisis Zones using YOLOv8

Nkamgan Nsonwang Laurent Stanix[a,*], Raissa Onanena Guelan[b], Mbous Ikong Jacques[b], Olivier Videme Bossou[b], Essimbi Zobo Bernard[a].

[a]*Department of Physics, University of Yaounde I, Yaounde, Cameroon*

[b]National Advanced School of Engineering, University of Yaounde I, Yaounde, Cameroon

ABSTRACT

Enhancing security measures in the crisis-affected regions of Cameroon, which have faced prolonged unrest over the years, is critically important. Detecting close-combat weapon within dense crowds can significantly improve surveillance and safety amidst ongoing challenges. These regions require advanced security solutions to address persistent threats faced by the local population. This study develops a detection model based on the You Only Look Once (YOLOv8) architecture to accurately identify and segment sticks and machetes in these crisis-affected areas. By assembling a diverse dataset that captures various scenarios, orientations, and lighting conditions, the model learns to recognize the distinctive features of these objects. By enhancing security measures through advanced technology, this study aims to contribute to ongoing efforts to safeguard communities and restore stability in these troubled areas.

Keywords: Object Detection, YOLOv8, Crisis-Affected Regions, *Cameroon Security.*

I. INTRODUCTION

Cameroon is currently facing insecurity challenges, characterized by frequent attacks and protests involving weapons such as machetes and sticks. In major cities across the country, numerous surveillance cameras have been installed, with personnel constantly monitoring video feeds. However, the proposed model aims to automate this monitoring process by detecting these threatening objects in real-time. Detecting objects like sticks and machetes in crowded environments presents a unique challenge that demands a precise and efficient detection system.

There has been great advancement in the field of object detection lately. Of the numerous object detection approaches, we can break the journey into the pre-2012 era or pre AlexNet era and the post-2012 era. The pre-2012 era includes multiple object detection algorithms such as Histogram of Oriented Gradients (HOG), Haar cascades, some variations of Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), etc. The post-2012 era includes Regions with Convolutional Neural Networks (RCNN), Fast RCNN, Faster RCNN, YOLO, Single Shot Detector, etc. In recent years, significant advancements have been made in artificial intelligence technology, particularly in the field of machine vision. These breakthroughs have led to the development of various neural network models that provide high accuracy and rapid response times. This progress offers a new solution for object detection, significantly reducing the need for human and material resources while improving detection accuracy and efficiency. The You Only Look Once (YOLO) model series is a widely used target detection framework that has been extensively applied to detect defects, achieving commendable accuracy and detection results. Similarly, the Region-based Convolutional Neural Networks (Faster R-CNN) model is another popular target detection framework that uses candidate region extraction and classification regression networks to accurately locate and identify objects.

Additionally, several studies have combined deep learning models with image segmentation techniques to enable precise segmentation and detection. Notable examples include the use of models like U-Net and Mask R-CNN for localizing and segmenting defective regions [1].

Some regions in Cameroon have been plagued by a longstanding crisis, leading to increased violence and threats to human lives. In such volatile environments, the ability to detect and identify potential weapons, such as sticks and machetes, becomes crucial for ensuring the safety and security of individuals. Traditional methods of manual inspection and surveillance is time-consuming, resource-intensive, and often prone to errors. Therefore, there is a pressing need for an efficient and accurate computer vision model that can automatically detect these weapons in real-time. In this research article, we propose the utilization of the YOLOv8 model, a state-of-the-art object detection algorithm, to address the aforementioned challenge. YOLOv8 (You Only Look Once version 8) is renowned for its exceptional speed and accuracy, making it an ideal candidate for real-time applications. By training the YOLOv8 model on a custom dataset specifically curated for crisis affected regions of Cameroon, we aim to develop a solution that can effectively identify sticks and machetes, enabling timely intervention and prevention of potential violence. By successfully implementing the YOLOv8 model for stick and machete detection, this research aims to contribute to the development of computer vision solutions that can enhance the safety and security of crisis-affected regions. The findings of this study have the potential to inform policymakers, law enforcement agencies, and humanitarian organizations, enabling them to take proactive measures and mitigate the risks associated with weapon-related violence. The

* laurent.nkamgan@facsciences-uy1.cm

Presented at the 2025 OkIP International Conference on Artificial Intelligence Frontiers (CAIF) in Oklahoma City, OK, USA, and Online, on April 2, 2025

YOLO series of algorithms is one of the fastest growing and best algorithms so far, especially the novel YOLOv8 algorithm released in 2023. YOLO is currently the most popular real-time object detector and is designed to combine the advantages of many real-time object detectors [2].

II. PROPOSED INNOVATION

Our contributions can be summarized as the following three points:

i) The proposed solution significantly enhances the security landscape in Cameroon by providing an automated and efficient method for detecting potential weapons like sicks and machetes in real-time, thereby addressing a critical gap in current security measures.

ii) Existing weapon detection systems primarily focus on conventional firearms and do not adequately account for the detection of improvised weapons such as sticks and machetes, which are prevalent in the context of the crisis-affected areas. By incorporating sticks and machetes detection into our model, we broaden the scope of weapon recognition.

iii) We have created and labelled a comprehensive dataset of sticks and machetes using Roboflow, providing a valuable resource for future research in object detection and violence prevention.

III. LITERATURE REVIEW

CNN, or Convolutional Neural Network, is a widely used machine learning technique in vision-related applications. It consists of several key processes: convolution, pooling, activation functions, and fully connected layers. Convolution acts as a feature extractor, capturing important patterns and features from the input data. Pooling is used for down-sampling, reducing the spatial dimensions of the data while retaining important information. Activation functions introduce nonlinearity to the algorithm, allowing it to handle complex features and patterns effectively. Fully connected layers are responsible for classifying the extracted features. They take the output from the previous layers and map it to the desired output classes. To train a CNN, a training dataset is required. The backpropagation technique is commonly used to train the algorithm by adjusting the weights and biases to minimize the classification error. The architecture of a CNN can vary depending on the specific application and the available training data. Different architectures have been introduced in recent years to improve performance and accuracy in various tasks.

Sanam et al. [3] developed in 2021 a method based on an integrated framework for reconnaissance security that aims to identify and distinguish weapons progressively. The proposed model utilizes computer vision techniques to detect unsafe weapons and alert security personnel in real-time.

Jamilu S et al. [4] in 2022 implemented YOLOv6, a single stage object detection framework, for industrial applications. The YOLOv6 architecture is inspired by the original YOLO architecture and offers improved detection accuracy and inference speed. The article compares YOLOv6 with previous versions of YOLO and demonstrates its superior performance in terms of mean average precision (mAP) and faster inference time. The proposed pipeline for weapon detection includes preprocessing of video frames, YOLOv6 object detection, GPS location tracking, database storage, and an alert system.

In August 2023, Liyao Lu, et al. [5] worked on developing the YOLOv8 algorithm architecture, which consists of Backbone, Neck, and Head components. The Backbone structure of YOLOv8 utilizes the concept of cross-level components (CSP) for feature fusion. The Neck section incorporates multi-scale feature fusion and separates the classification and detection heads. The improved model CSS (CSP-Swin-SoftNMS) is introduced, which combines CBAM (Channel Attention Module), Swin Transformer, Soft NMS (Non-Maximum Suppression), and Mosaic data augmentation. CBAM, the attention module reduces computation and parameters by decomposing the attention mechanism into channel attention and spatial attention. Swin Transformer enhances the model's ability to extract global features by using a sliding window multi-head self-attention model. Soft NMS replaces the traditional NMS algorithm with a Gaussian reset method to avoid missed detections in overlapping areas. Mosaic data augmentation randomly crops and stitches four images together to enrich the background and increase batch size during training.

Armstrong Aboah et al. [6] in 2023 conducted experiments on three single-stage object detection models: YOLOv5, YOLOv7, and YOLOv8.YOLOv5 Features a Backbone (CSPDarknet53 with 29 convolutional layers), Neck (SPP block and PANet), and Head for predicting object classes and bounding boxes. YOLOv7 Introduces compound scaling, Extended Efficient Layer Aggregation Network (EELAN), planned and reparametrized convolution, model re parameterization, and dynamic label assignment for improved learning and optimization. YOLOv8 Known for joint detection and segmentation, with a new architecture, improved convolutional layers, an anchor-free detection head, and feature pyramid networks. It uses the Darknet-53 backbone and offers a user-friendly API.The models were trained with optimal hyperparameters generated by a genetic algorithm, for 400 epochs with a batch size of 16 and an image size of 832x832. Test Time Augmentation (TTA) was used to enhance prediction accuracy.

Still in 2023 Nilesh et al. [7] propose a semantic segmentation architecture combining MobileNet and YOLOv8. MobileNet is chosen for its balance of speed and accuracy on mobile platforms, using depth wise and pointwise convolutions to reduce operations and parameters. Bottleneck layers and leftover connections enhance backpropagation and computation. YOLOv8 is compatible with earlier YOLO versions, facilitating evaluation and switching. The model is tested on the Mask Dataset, detecting persons with masks, without masks, and with improper masks. The system uses a 448x448 pixel input and a cost function combining class-specific scores and bounding box predictions, considering object scale and class scale. Evaluation criteria include overall

precision and recall. Datasets used are the Mask Dataset, ImageNet, and FDDB.

Not all authors face the challenges of precision and execution speed. With some prioritizing high precision, others focus on execution speed, creating a constant balancing act between these two metrics. For our project, we opted for the YOLOv8 model, which effectively balances both precision and execution speed, ensuring optimal performance without compromising either aspect. After evaluating several models, we found YOLOv8 to be superior in speed and accuracy, which is crucial for processing the substantial amount of data generated by continuous surveillance videos recorded day and night.

IV. METHODS

To enhance the robustness of our model, we curated a diverse set of images of sticks and machetes. These images were sourced from online repositories focusing on crisis-affected regions, as well as through custom photography sessions. The collected dataset includes variations in lighting conditions, backgrounds, and object orientations, ensuring that the model can effectively generalize across different scenarios.

For the annotation process, we utilized Roboflow, a powerful and user-friendly tool designed for computer vision tasks. we have allocated 86% of the dataset for training, 7% for validation, and 7% for testing.

Model Selection

In this work, we utilized the YOLOv8s segmentation model, which is a powerful and widely-used model for object detection and segmentation tasks. One of the key advantages of using the YOLOv8s model is its ability to leverage transfer learning. Transfer learning is a technique where a pre-trained model, which has been trained on a large and diverse dataset, is fine-tuned on a smaller, more specific dataset. In this exercise, we fine-tuned the pre-trained YOLOv8s model on our annotated dataset of sticks and machetes. By leveraging transfer learning, we take advantage of the knowledge and features that the pre-trained model has already learned from its extensive training on a diverse range of objects. This significantly reduces the amount of data and computational resources required to train our model from scratch. Also, transfer learning improves the performance and accuracy of our model on our specific task of detecting and segmenting sticks and machetes. The pre-trained model has already learned to recognize various object features, such as edges, textures, and shapes, which are relevant for our task. By fine-tuning the model on our dataset, it can quickly adapt and specialize in detecting and segmenting sticks and machetes with higher accuracy.

Training Process

The YOLOv8 model was implemented and trained using the PyTorch framework. The experiments were conducted on a 64-bit Windows 11 operating system within SageMaker Studio Lab. The model was trained for 100 epochs with a learning rate of 0.1 and batch size of 16. We trained the model for stick detection and segmentation separately from machete detection and segmentation because our dataset was collected separately for sticks and machetes. In this light, we trained two separate models as it is uncommon for people to protest using both sticks and machetes simultaneously. Typically, individuals choose one tool over the other when participating in protests. However, results of the training in the exercise can be combined to create a unified model.

Evaluation Metric

In this work, the model's performance is evaluated using the Mean Average Precision (mAP) metric. The mAP is calculated by taking the average of the Average Precision (AP) values for each target class.

Future Work

While our current model shows promising results, there are several avenues for future work that could further enhance its robustness and performance.

1) Collecting More Data
2) Training Using Other Hyperparameters
3) Increasing the Number of Epochs While Monitoring Overfitting:
4) Data Augmentation
5) Model Architecture Tuning
6) Cross-Validation

By applying these strategies, we can potentially achieve a better overall performance for our model.

REFERENCE

[1] Yan, W; Kehua, Z.; Ling, W; Lintong, W. Improved YOLOv8 Algorithm for Rail Surface Defect Detection, in IEEE Access, vol. 12, pp. 44984- 44997, 2024.

[2] Lou, H.; Duan, X.; Guo, J.; Liu, H.; Gu, J.; Bi, L.; Chen, H. DCYOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor. Electronics 2023, 12, 2323.

[3] Narejo, S.; Pandey, B.; Esenarro, V.; Rodriguez, C.; Anjum, M. Weapon Detection Using YOLO V3 for Smart Surveillance System. Mathematical Problems in Engineering. 2021. 1-9. 10.1155/2021/9975700.

[4] Jamilu, S.; Majeed, H. Weapon Detection for Smart Surveillance System using YOLOV6. Global Journal of Research in Humanities and Cultural Studies ISSN: 2583-2670 (Online) Volume 02— Issue 05 — Sept.-Oct. 2022

[5] Liyao, L. Improved YOLOv8 Detection Algorithm in X-ray Contraband. Advances in Artificial Intelligence and Machine Learning. 2023:3(3):72.

[6] Armstrong, A.; Bin, W.; Ulas, B.; Yaw, A. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Vancouver, BC, Canada,2023, pp.5350-5358, Doi: 10.1109/CVPRW59228.2023.00564.

[7] Motwani, N.; Soumya, S. Human Activities Detection using Deep Learning Technique- YOLOv8. ITM Web of Conferences. 56. 10.1051/itmconf/ 20235603003.