

The MIT logo is displayed in a bold, dark blue, sans-serif font.

Academy of
Engineering

(An Autonomous Institute Affiliated to Savitribai Phule Pune University)

AUTOMATED WEAPON DETECTION AND CLASSIFICATION

Using YOLOv12, EfficientNetB0 &
Grad-CAM

Mayur Kapgate - 202201040065
Anirudha Gapat - 202201040067
Ganesh Atre - 202201040068



CONTENTS

- Motivation & Objectives
- System Architecture
- Dataset & Preprocessing
- Model Training
- Explainability with Grad-CAM
- Results & Performance
- Use Cases & Applications
- Limitations & Future Work



MOTIVATION & OBJECTIVES

- Need for real-time, scalable threat detection
- Manual methods are slow & error-prone

Objective: End-to-end pipeline to detect and classify weapons

- Provide explainable outputs for user trust



SYSTEM ARCHITECTURE

- YOLOv12 for detection
- EfficientNetB0 for classification
- Grad-CAM for explainability
- Pipeline: Image → Detection → Crop → Classification
→ Grad-CAM



DATASET & PREPROCESSING

- Custom dataset: 700+ train, 140 val, 9 classes
- Resize to 480×480, normalize, augment (flips, mosaic, color jitter)
- Split: 70% train, 20% val, 10% test



MODEL TRAINING

- Detection: YOLOv12x, epochs=70, SGD, aug: flip, mosaic, HSV jitter
- Classification: EfficientNetB0, epochs=10, Adam, LR=1e-4, batch=32

DETECTION PERFORMANCE

Metric	Value
Precision	0.860
Recall	0.678
F1 Score	0.758
mAP@0.5	0.815
mAP@0.5:0.95	0.656

DETECTION PERFORMANCE

Model	mAP@0.5	F1 Score	Params (M)	Inference (ms)
ViT	0.78	0.82	86	40
ResNet50	0.80	0.83	25	30
EfficientNetB0	0.815	0.83	5.3	27



USE CASES & APPLICATIONS

- Public Safety & Law Enforcement
- Smart City CCTV Integration
- Access Control
- Forensics & Threat Analysis
- Event Security



CONCLUSION

Our system combines the real-time detection power of YOLOv12 with the efficient feature-extraction capabilities of EfficientNetB0 and the interpretability of Grad-CAM into a unified pipeline. This architecture delivers strong performance –achieving high mAP and F1 scores on diverse weapon classes –while maintaining low inference latency suitable for live monitoring. By overlaying Grad-CAM heatmaps on each detection, the solution remains both scalable and transparent, giving operators clear insight into the model’s decision-making process.



THANK YOU