IT469 AIH - Project Evaluation [Jul-Nov 2024]

# Optimized YOLO11 for Fallen Person Detection


**Group Details**

| | |
|---|---|
| Raghavendra | (211AI029) |
| Venkatesh BM | (211AI039) |
| Vinayaka SN | (211AI040) |
| Vivek Vittal Biragoni | (211AI041) |

# INTRODUCTION

- Falls pose significant risks of injury among the elderly, disabled children, and patients with neurological or cardiovascular diseases.

- Existing methods, such as wearable sensors, are often intrusive and uncomfortable for users.

- Current computer vision-based fall detection systems struggle with real-time performance, especially in complex environments.

- The design of YOLOv8 is optimized in this study to differentiate between "fall" and "no-fall" occurrences.

# Relevant Research works

- Habib Khan et al. developed a fall detection system using YOLOv8, enhanced with a Focus Module and CBAM, leveraging the DiverseFALL10500 dataset for improved accuracy and real-time processing in varied conditions, making it suitable for healthcare applications.

- Zhong Zhang et al. introduced an occluded fall detection benchmark dataset using Kinect depth cameras, evaluating four depth-based methods and revealing performance degradation in occluded scenarios compared to non-occluded ones.

# Relevant Research works

- Zhigang Yu et al. proposed a fall detection system called Fed-ELM, utilizing Federated Learning and Extreme Learning Machine for enhanced privacy and continuous learning, achieving improved accuracy in distinguishing falls from daily activities while adapting to individual movement patterns.

- Yu et al. 2022 uses a combination of Online Extreme Learning Machine (OS-ELM) and Federated Learning (FL)to tackle the problem of detecting falls in older people. In contrast to conventional methods that employ data from youth, which results in lower accuracy because of variations in movement patterns

# Outcome of Literature Review

- The first paper (Habib Khan et al., 2024) enhances YOLOv8 with CBAMs and a focus module, achieving superior real-time fall detection with high mAP, precision, recall, and F1-score.

- It provides a large, annotated dataset of 10,500 images, capturing diverse lighting, angles, and settings, improving model generalization and robustness.

- Public availability of code and dataset makes it ideal for real-time vision-based FD, addressing gaps like inefficiency and lack of scalable datasets in existing systems.

# Problem Statement and Objectives

Existing fall detection methods are often intrusive and ineffective. Our project aims to develop a real-time, accurate, non-intrusive and user-friendly fall detection system using an enhanced version of YOLOv8.

**Objectives**:

- Create a system that can detect falls in real-time from both live video feeds and locally uploaded videos.

- Integrate advanced techniques, such as convolutional block attention modules (CBAMs), with YOLOv8 or better to improve the accuracy of fall detection in complex environments.

- Implement a dashboard that provides an intuitive interface for users to view detection results and manage video uploads.

# Existing METHODOLOGY

- Existing Methodology: Insights from "Visionary Vigilance"

- Utilized YOLOv8 for feature extraction in fall detection tasks.

- Incorporated Convolutional Block Attention Module (CBAM) for enhanced attention:

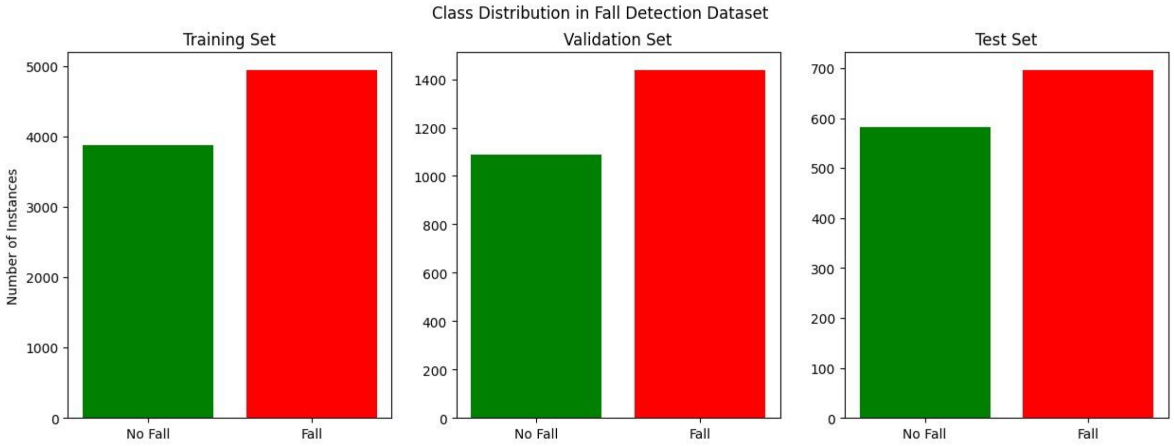
    - **Spatial Attention:** Reduced background noise, emphasized key fall areas.

    - **Channel Attention:** Focused on relevant body orientation and posture features.

- Focused on general fall detection datasets for validation and training.

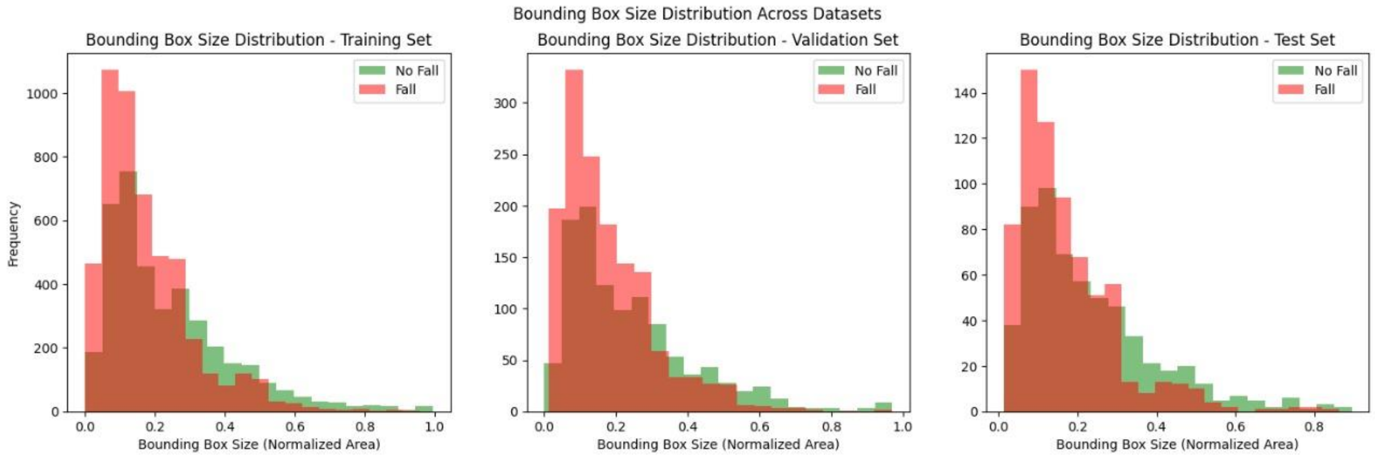
# Dataset Overview

- About 10,500 annotated samples, including both "Fall"and "No Fall" cases, make up the dataset used for fall detection It provides a large, annotated dataset of 10,500 images, capturing diverse lighting, angles, and settings, improving model generalization and robustness.



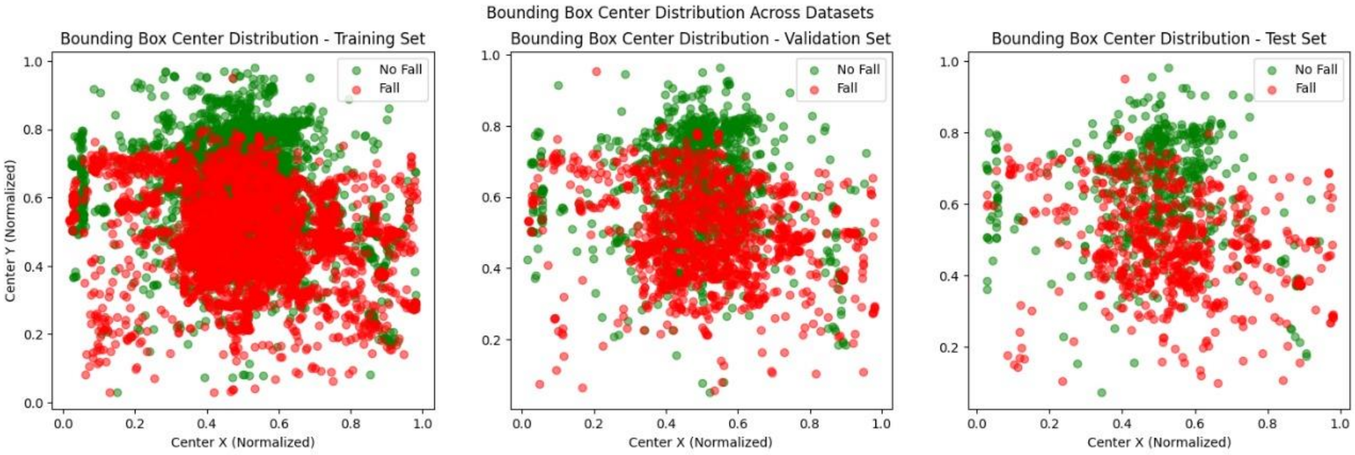
# Dataset Overview

- An evenly distributed dataset is reflected in this uniform distribution across subsets
- The overlap in size distributions between the "Fall" and "No Fall" groups suggests size might not be a key differentiator.
- This emphasizes how sophisticated feature ex-traction methods are required to reliably distinguish between the two classes.



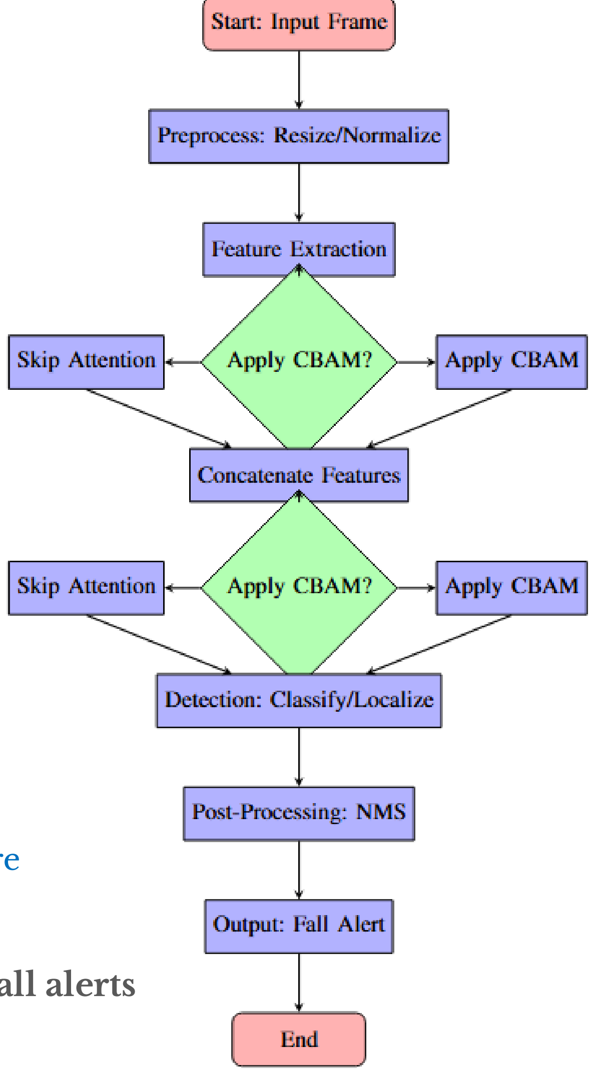
# Dataset Overview

- The features of the dataset are further clarified by the bounding box center distribution
- The Training, Validation, and Test subsets all exhibit this spatial bias, which guarantees that the dataset accurately depicts actual fall detection situations.



# PROPOSED ENHANCEMENTS

- Enhancing YOLO11 with CBAM.
- Used DiverseFALL10500 dataset for training and validation.
- Backbone Modifications:
  - CBAM added after intermediate bottlenecks (P3/8).
  - CBAM integrated before deeper convolutions (P5/32).
- Head Modifications:
  - CBAM before upsampling to preserve small-scale details.
  - CBAM placed at P4/16 and P5/32 scales for large and medium feature enhancement.
  - User-friendly interface for video upload and results display. **Fall alerts sent via Telegram**. Notifications are non-redundant.



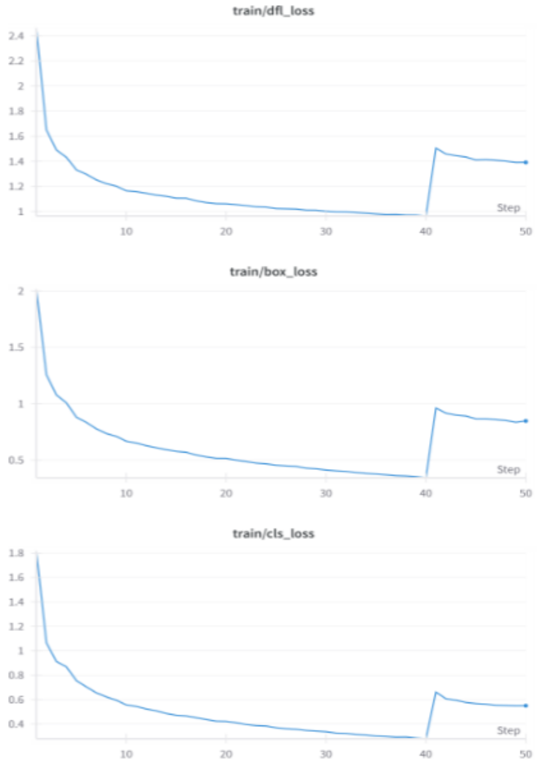
# WORK DONE

- Enhanced the model's learning process using advanced loss metrics:
  - Distribution Focal Loss (DFL) for improved bounding box localization.
  - Box Loss for accurate object boundary predictions.
  - Classification Loss for precise object labeling.
- Validation losses analyzed to ensure generalization and minimize overfitting.
- Implemented dynamic learning rate schedules for stable convergence.

# WORK DONE

Training loss analysis:

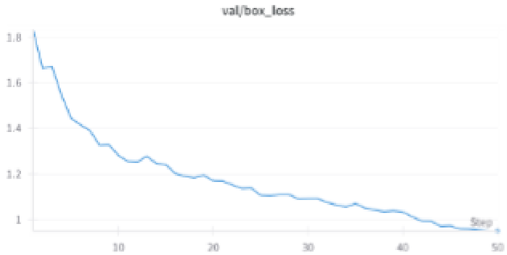
- DFL Loss:
    - Improves bounding box accuracy for detecting critical areas.
    - Steady reduction indicates enhanced localization precision.
- Box Loss:
    - Ensures better alignment with ground truth bounding boxes.
    - Essential for accurate boundary detection in healthcare.
- Classification Loss:
    - Reduces false positives and negatives, ensuring high confidence in object categorization.
    - Vital for applications like fall detection.

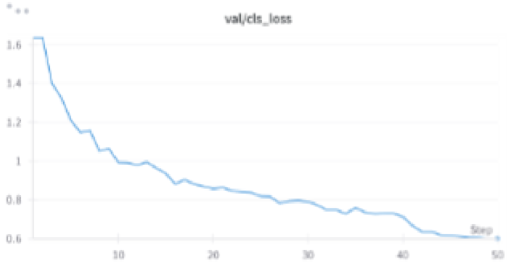


# WORK DONE

## Validation loss analysis

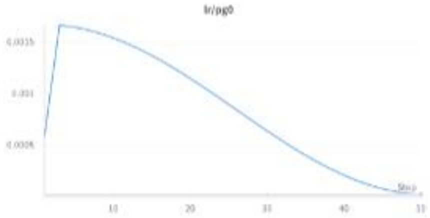
- Validation loss trends demonstrate strong generalization.
  - **DFL Loss:** Accurate bounding box localization on unseen data.
  - **Box Loss:** Reliable object boundary predictions across varied conditions.
  - **Classification Loss:** High confidence in categorizing unseen data.
- Balanced training and validation losses indicate minimal overfitting.

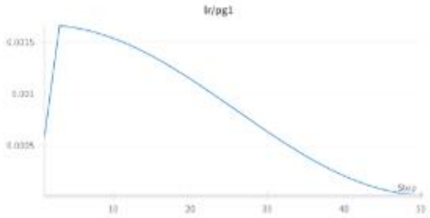



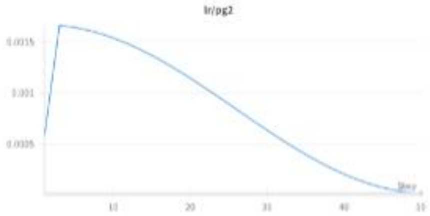



# WORK DONE

## Dynamic Learning Rate Optimization




- **Cosine Annealing Schedule:**
    - Ensures fast convergence early and stability later.
- **Warmup Phase:**
    - Gradual increase in learning rate prevents instability at the start.
- **Parameter-Specific Learning Rates:**
    - **pg0 (Backbone):** Stable adjustments for feature extraction.
    - **pg1 (Neck):** Fine-tuning feature aggregation.
    - **pg2 (Head):** Dynamic rates for precise classification and detection.

# WORK DONE

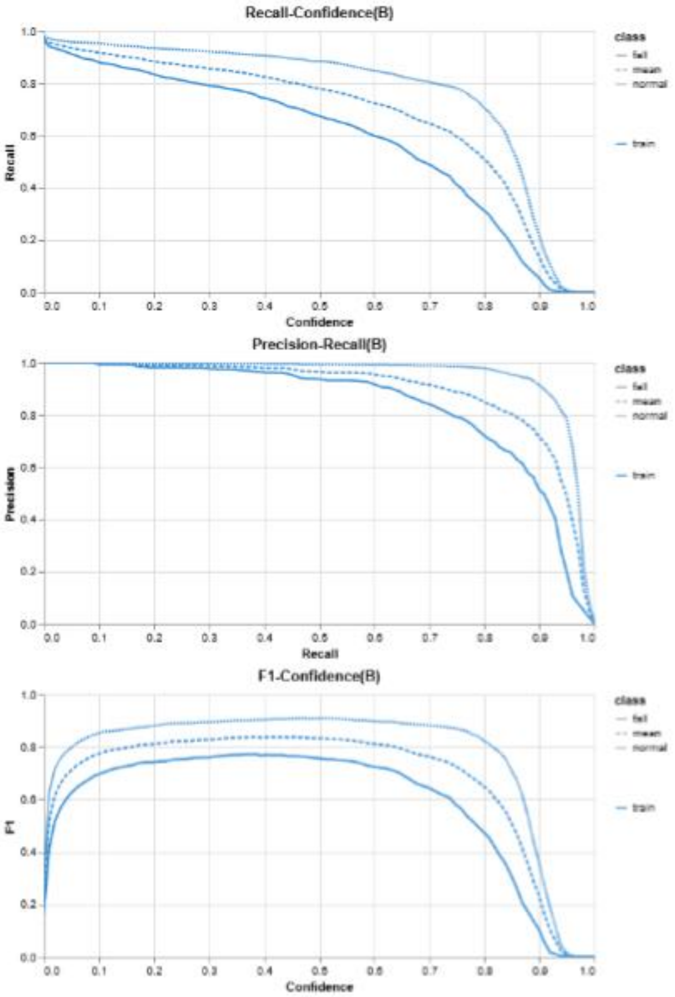
## Performance Metrics and Batch Inference Results

- **Precision-Recall Analysis:** $F1 = 2 \times \left( \frac{\mathrm{Pe} \times \mathrm{Re}}{\mathrm{Pe} + \mathrm{Re}} \right)$
  - High recall ensures minimal missed detections in healthcare applications.

$$AP = \sum_{i=1}^{c} \mathrm{Pr}(j) \times \Delta\ \mathrm{Re}(j)$$

- **mAP50-95 and mAP50:** $mAP = \frac{1}{2} \sum_{i=1}^{n} AP_i$
  - Comprehensive evaluation of detection accuracy over varying IoU thresholds.

- **Batch Inference:**
  - Accurate detection and classification on training batch images.



# RESULTS

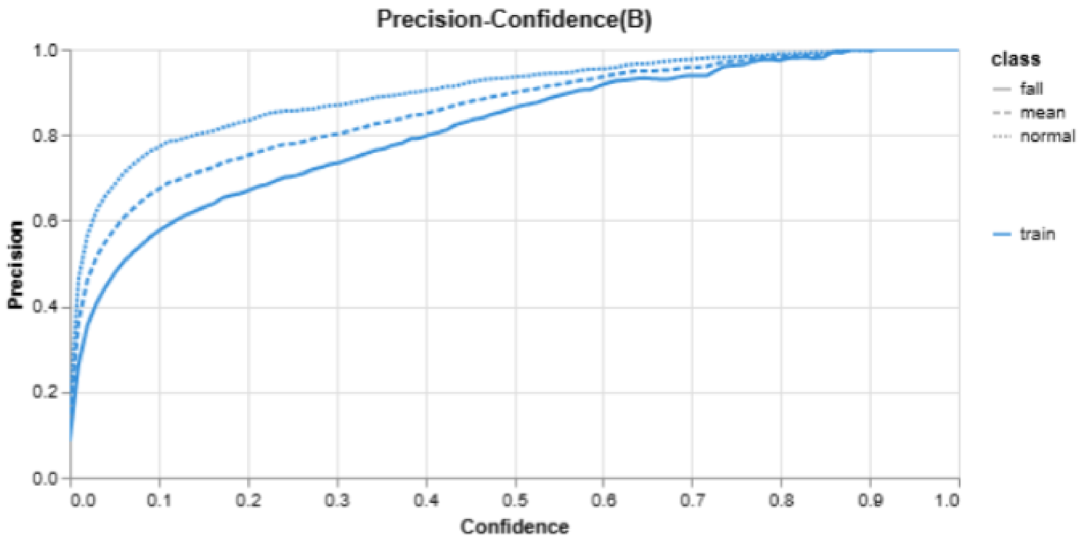


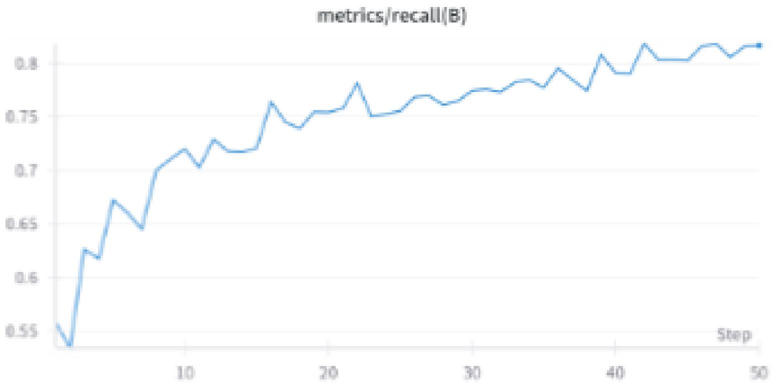

Fig. 8: Precision-Confidence Curve

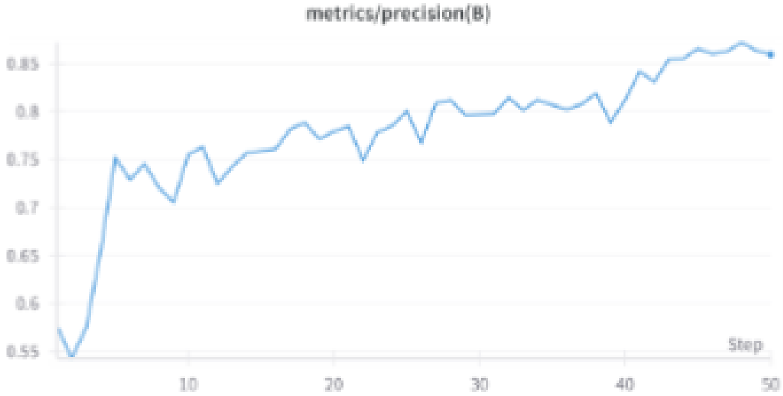
The precision-confidence chart provides insight into classification reliability in healthcare applications by displaying model confidence across precision scores.

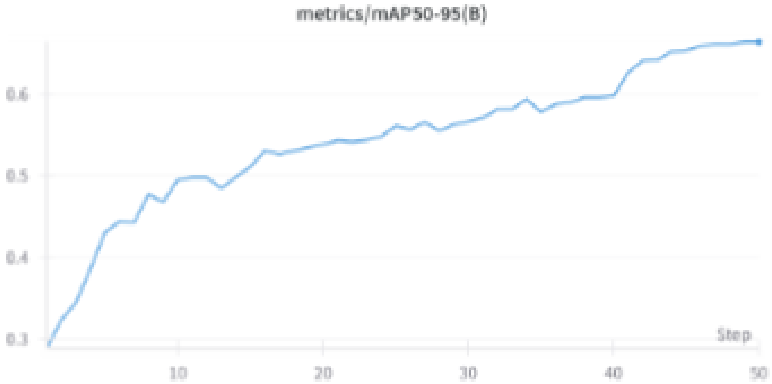
# RESULTS

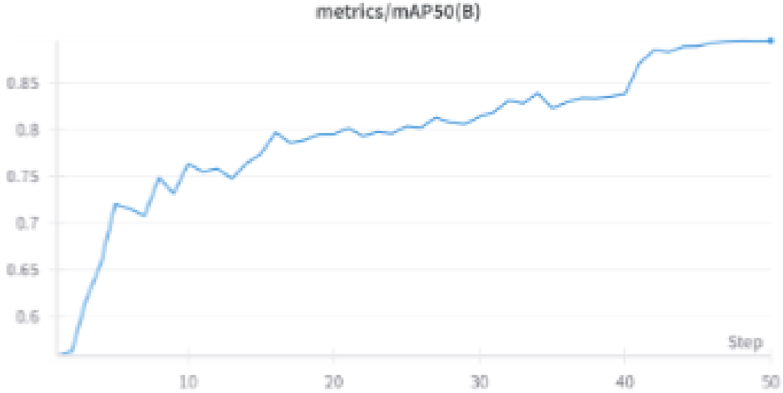
$$\mathrm{Re} = \frac{TP}{TP+FN}$$

$$\mathrm{Pr} = \frac{TP}{TP+FP}$$



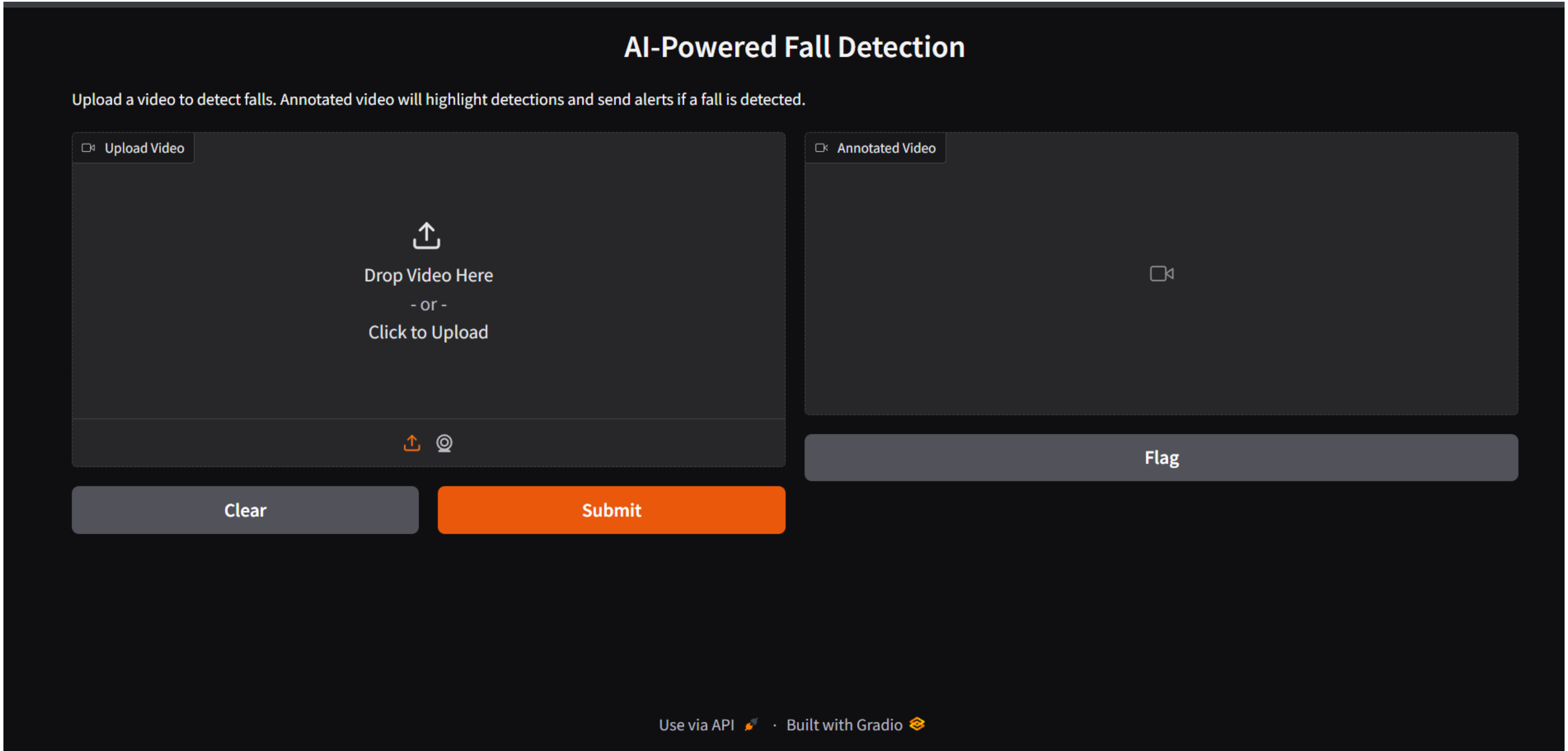



# Comparison with other existing works

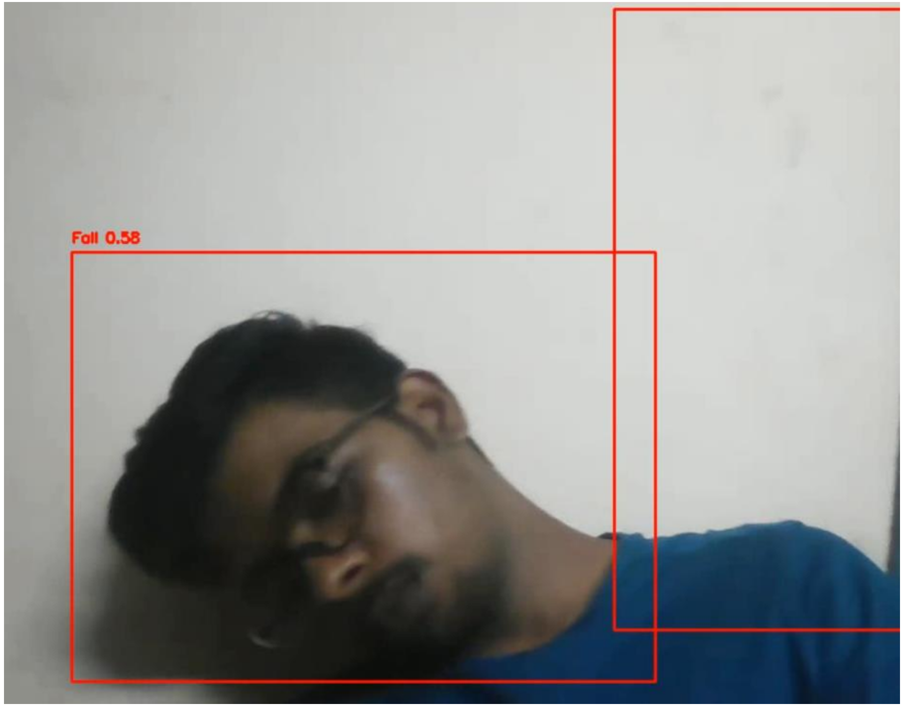
| S No. | Model | mAP | Precision | F1-Score | Recall | Inf. Time (ms) |
|---|---|---|---|---|---|---|
| 1 | YOLOv8M | 0.899 | 0.880 | 0.868 | 0.851 | 12.3 |
| 2 | Base Paper | 0.935 | 0.900 | 0.884 | 0.859 | 10.5 |
| 3 | Our with 50 Epochs only | 0.91 | 0.897 | 0.879 | 0.857 | **9.8** |

Note: Our results are based on 50 epochs of training, whereas the benchmark models were trained for 100 epochs.

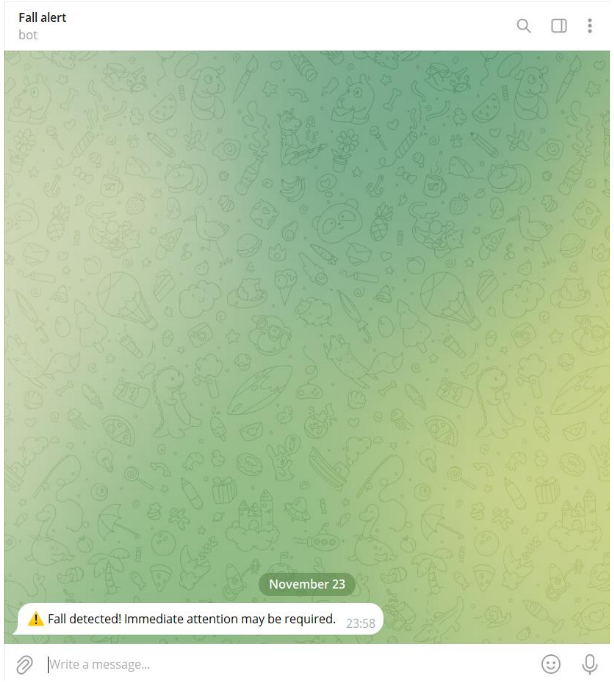
# RESULTS


AI-Powered Fall Detection
Upload a video to detect falls. Annotated video will highlight detections and send alerts if a fall is detected.
Upload Video
Drop Video Here
- or -
Click to Upload
Annotated Video
Flag
Clear
Submit
Use via API
Built with Gradio

# RESULTS




Detection of fall of a person from the input video given to the web interface




Telegram notification alert

# CONCLUSION AND FUTURE WORK

- Optimized YOLO11 with CBAM enhances fall detection accuracy in diverse environments.
- Deployment via Gradio with telegram ensures real-time alerts and practical usability.
- Results demonstrate the model's reliability and scalability for dynamic healthcare applications.
- Incorporate advanced mechanisms to handle occlusions and crowded environments.
- Optimize the system for deployment on edge devices.
    - Export to ONNX or OpenVINO for up to 3x CPU speedup.
    - Export to TensorRT for up to 5x GPU speedup.

# Individual Contribution

- Vivek:
  - Trained custom modified architecture of yolo11 for FD.
- Vinayaka:
  - Developed User interface and integrated telegram alert notifications when fall is detected.
- Venkatesh:
  - Worked on existing research methods and model selection(CBAM).
- Raghavendra:
  - Worked on DiverseFallen10500 dataset and visualization methods.

# REFERENCES


- Khan, H., Ullah, I., Shabaz, M., Omer, M. F., Usman, M. T.,Guellil, M. S., and Koo, J. (2024). Visionary vigilance: Optimized yolov8 for fallen person detection with large-scale benchmark dataset Image and Vision Computing Accessed: 2024-06-18
- Alam, E., Sufian, A., Dutta, P., and Leo, M. (2022). Vision-based human fall detection systems using deep learning: A review. Computers in Biology and Medicine Accessed: 2024-06-18
- Yu, Z., Liu, J., Yang, M., Cheng, Y., Hu, J., and Li, X. (2022) An elderly fall detection method based on federated learning and extreme learning machine (fed-elm). IEEE Access. Accessed: 2024-06-18