

Project Title: Animal Detection, Identification, and Counting from Video

1. Objective

The objective of this project is to develop a computer vision system capable of detecting animals in wildlife video footage, identifying their species, tracking individual animals across frames, and generating accurate counting statistics. The system will process pre-recorded wildlife videos and produce structured outputs suitable for analysis and reporting.

2. Target Species Scope

The system will detect and classify the following animal species in the initial implementation phase:

- Elephant
- Deer
- Tiger

The system will be designed to allow future extension to additional species through retraining or fine-tuning.

3. Definition of Animal Identification

Animal identification is defined as the process of:

- **Detecting the presence of an animal** in a video frame using a bounding box, and
- **Assigning the correct species label** to the detected animal using a trained object detection model.

Each detection will include the following attributes:

- Bounding box coordinates
- Species label
- Confidence score
- Frame number

Identification will be performed at the species level.

4. Definition of Animal Counting

The system will support two types of counting:

4.1 Per-Frame Counting

Per-frame counting refers to counting the number of animals of each species present in a single video frame.

Example:

Frame 001:

Elephant: 3

Deer: 2

This count is based on detections within that specific frame.

4.2 Unique Per-Video Counting

Unique per-video counting refers to counting the **number of distinct individual animals appearing throughout the entire video.**

This will be achieved using multi-object tracking, where **each detected animal** is assigned a **unique tracking ID**. The system will count unique IDs per species to avoid double counting the same animal across multiple frames.

Example:

Video Summary:

Elephant: 5 unique individuals

Deer: 8 unique individuals

5. Tracking Requirement

The system will implement object tracking to:

- Assign persistent unique IDs to animals
- Maintain identity consistency across frames
- Prevent duplicate counting of the same animal

Tracking will operate at the frame sequence level within each video.

6. Expected Output Formats

The system will generate structured outputs in CSV and JSON formats.

6.1 CSV Output (Frame-wise counts)

Example format:

frame_number, species, count

001, elephant, 3

002, deer, 2

6.2 JSON Output (Video summary)

Example format:

```
{  
  "video_name": "sample_video.mp4",  
  "unique_counts": {  
    "elephant": 5,  
    "deer": 8,  
    "monkey": 3  
  }  
}
```

7. System Inputs and Outputs

Input

- Wildlife video files (MP4, AVI, or similar formats)

Outputs

- Frame-wise species count (CSV)
- Unique per-video species count (JSON / CSV)
- Annotated video with bounding boxes and tracking IDs

8. Assumptions and Constraints

- Videos contain visible animals within detectable resolution.
- Species classification is limited to the defined target species list.
- Tracking identity consistency is maintained within individual video clips.
- Severe occlusions or poor visibility may affect detection accuracy.