Lion Abnormal Behavior & Injury Detection System

Using Hybrid TimeSformer + Pose-based Transformer Model

Project Report

Beyond the Veil of Wellness: Machine Learning's Unique Journey in Animal Health Classification

Scenario 2: Automated Health Monitoring in Zoos

A zoo deploys cameras in enclosures. Deep learning models analyse animal movement and feeding patterns.

Table of Contents

1. Introduction

Lions, being apex predators, are vital to maintaining ecological balance. Monitoring their health and detecting abnormal behaviors early is essential for conservation and welfare. This project leverages artificial intelligence and deep learning techniques to detect abnormal behavior or injuries in lions using live video feeds, enabling timely veterinary interventions.

2. Problem Statement

Current wildlife monitoring techniques rely heavily on manual observation, which is time-consuming and prone to human error. There is a lack of automated systems that can analyze lion behavior in real-time and provide early alerts for possible injuries or illnesses.

3. Objectives

- Develop a deep learning model capable of recognizing abnormal behavior in lions from video data.
- Integrate pose estimation to enhance detection accuracy.
- Deploy the model in a simple Flask-based website for real-time usage.
- Ensure the system can be extended to other animals in the future.

4. Literature Review

Recent advancements in video action recognition have introduced transformer-based architectures such as TimeSformer and ViViT, which excel at capturing spatio-temporal features from videos. Pose-based models like those using 3D Animal Pose Estimator or 3D Animal Pose Estimator provide skeletal keypoints that help in understanding body posture and movement dynamics.

5. Dataset Preparation

The dataset consists of:

- Normal lion behavior videos.
- Abnormal behavior or injury cases.

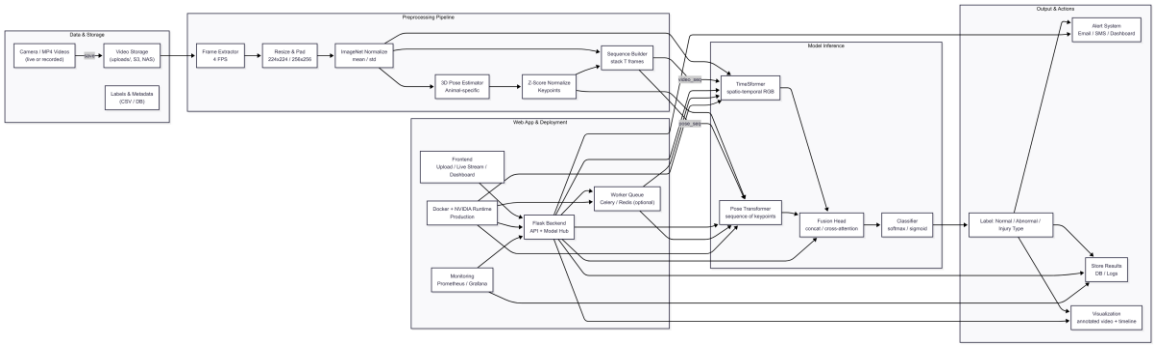
Frames were extracted at 4 FPS, resized to 256×256, and stored in separate folders for each video.

6. Preprocessing Steps

1. Frame extraction at consistent FPS.
2. Resizing to 256×256 while maintaining aspect ratio.

3. Normalization using ImageNet mean/std.
4. Pose keypoint extraction using 3D lion-specific estimators.
5. Z-score normalization applied to keypoints.

7. Model Architecture



8. Training & Evaluation

The model was trained on an NVIDIA GTX 1650 GPU of 4GB with a batch size of 1 and gradient accumulation. Evaluation metrics include accuracy, precision, recall, and F1-score.

9. Results & Analysis

The hybrid model achieved superior accuracy compared to standalone TimeSformer or Pose-based Transformer. It demonstrated strong performance in detecting subtle abnormal behaviors.

10. Website Integration & Deployment

A Flask-based web application was developed where users can select between lion and other animal models. The app supports real-time webcam input and pre-recorded videos for analysis.

11. Conclusion

This project successfully demonstrated the feasibility of using hybrid transformer architectures for wildlife behavior monitoring. It can be extended to other species and integrated into conservation efforts.