In20-S5-CS3501 Data Science and Engineering Project
Department of Computer Science and Engineering
University of Moratuwa

REAL TIME HUMAN ACTIVITY DETECTION UNDER DARK LIGHTING CONDITIONS

Team Members:

Name: Vihidun D.P.T
Index: 200682T
Email: tharoosha.20@cse.mrt.ac.lk

Name: Umayanga S. A.I
Index: 200671J
Email: umayanga.20@cse.mrt.ac.lk

Name: Thusyanthan J.
Index: 200662H
Email: thusyanthan.20@cse.mrt.ac.lk

Group ID: 26
Project ID: 04
Mentor: Dr. Chathuranga Hettiarachchi

1. Executive Summary:

In this project, we aim to develop a real-time human activity detection system that can operate under dark lighting conditions. The system will use existing AI tools for human detection and action classification, allowing users to receive alerts when a human is detected on CCTV cameras installed in a premises. The system will also provide face identification capabilities. Additionally, the project involves engineering a scalable software platform and data processing workflow that enables easy integration of AI systems and provides alerts through a mobile app. The final system will be deployed on a free cloud server and accessible through a mobile app and a suitable web front-end for administration.

2. Problem Statement:

The project aims to develop a reliable real-time human activity detection system that can operate effectively in dark or night-vision lighting conditions. The need for such a system arises from the increasing demand for enhanced security and surveillance in various settings, such as modern houses, public spaces, and autonomous vehicles operating at night.

The key problems the project is aiming to solve are as follows:

1. Reliable Human Activity Detection: Existing human activity detection systems may struggle to perform well in low-light or dark environments, resulting in reduced accuracy and missed detections. The proposed system aims to address this issue by incorporating advanced algorithms and techniques to effectively detect human presence and activities even in challenging lighting conditions.
2. Action Classification: The project aims to go beyond simple human detection and also classify the actions or activities of detected individuals. This level of understanding can provide valuable context to the surveillance system, distinguishing between normal activities and potentially threatening behaviors.
3. Face Identification: In addition to detecting humans and their actions, the system also seeks to incorporate face identification capabilities. This feature allows for recognizing known individuals and potential intruders, providing an extra layer of security.

Importance and Impact:

1. Enhanced Security: The system will significantly enhance security in modern houses equipped with CCTV cameras, ensuring that any suspicious activities or intrusions are promptly detected and reported.
2. Night Surveillance: Night surveillance is a critical aspect of security, and the system will enable reliable monitoring and protection during nighttime hours when visibility is often reduced.
3. Reduced False Alarms: The incorporation of advanced algorithms for human activity classification and face identification can help reduce false alarms, minimizing unnecessary notifications and ensuring that alerts are relevant and reliable.

Overall, the project's success will positively impact the security and safety of individuals, households, and communities, making it an essential technological advancement with far-reaching implications.

3. Data Description:

The data used in the project is the Action Recognition in the Dark (ARID) dataset, specifically the updated version ARID (v1.5). The ARID dataset is designed to facilitate action recognition in low illumination conditions, making it suitable for scenarios such as night surveillance and self-driving at night. The dataset has been expanded and updated to provide a comprehensive set of video clips capturing various actions in dark environments.

The ARID (v1.5) Dataset, collected by researchers for action recognition in dark videos, consists of 5,572 unstructured video clips capturing various actions performed by individuals in low-light or dark conditions. It includes 11 action categories, recorded in 24 scenes (12 indoor, 12 outdoor), simulating challenging lighting conditions for evaluating action recognition performance. The dataset benefits from more than 15 volunteers, ensuring a diverse set of actions and variations for comprehensive research and is publicly available for research purposes.

The data collection process involved recording individuals performing various actions in both indoor and outdoor settings, under low-light conditions or in complete darkness. The collected videos were then categorized based on the actions performed and the environmental scenes.

4. Methods:

In the project, the main focus is on human action recognition in dark videos using the ARID dataset and the proposed Dark Temporal Consistency Model (DTCM) neural network architecture. The data science methods, techniques, and tools involved in the project can be outlined as follows

The ARID (v1.5) dataset contains video clips capturing human actions in low-light conditions. Data cleaning and preprocessing will be necessary to ensure the videos are in a consistent format and ready for model training. This may involve handling missing or corrupted data, resizing frames, and normalizing pixel values. Action recognition in videos is a computer vision task. Techniques such as **feature extraction**, **optical flow estimation**, and **motion analysis** may be employed to capture spatio-temporal patterns in the video clips and represent actions effectively. The proposed **Dark Temporal Consistency Model (DTCM)** is a novel neural network architecture specifically designed for action recognition in dark videos. Deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), will be utilized to build and train the DTCM. To leverage existing knowledge and improve model performance, transfer learning may be applied. Pretrained models on large-scale datasets for action recognition (e.g., ImageNet, Kinetics) can be used as a starting point and fine-tuned on the ARID dataset. PyTorch, a popular deep learning framework, will likely be used to implement and train the neural network models. Other libraries, such as NumPy, scikit-learn, and OpenCV, might also be used for data manipulation, evaluation, and visualization.

To gain insights into the model's decision-making process, techniques for model interpretation, such as **saliency maps** or **attention mechanisms**, may be employed. Experimentation with hyperparameter tuning, such as learning rates, batch sizes, and network architecture choices, will be conducted to find the optimal configuration for the DTCM. The project will involve a combination of data exploration, model development, training, and evaluation to achieve a robust and accurate human action recognition system for dark videos using the ARID dataset and the proposed DTCM architecture.

5. Expected Outcomes and Success Criteria:

The expected outcome of the project is a real-time human activity detection system that is capable of reliably operating under dark or low-lighting conditions. The system will be designed to accurately detect humans, classify their actions, and provide face identification for enhanced security and surveillance. The key goals of the project include achieving high accuracy in human detection and action classification, efficient resource utilization for real-time processing, and seamless integration of AI tools into a scalable software platform. Additionally, the project aims to provide a user-friendly mobile app and web front-end for easy accessibility and administration of the system.

The project will be considered successful if it meets the following criteria:

1. High Accuracy: The system should demonstrate a high level of accuracy in detecting humans, classifying their actions, and identifying faces. The accuracy metrics should be significantly improved compared to existing methods.
2. Real-time Operation: The system should be capable of processing video streams in real-time, providing timely and instantaneous results. Real-time operation is crucial for applications like security and surveillance.
3. Efficient Resource Utilization: The system should be resource-efficient, minimizing computational requirements and memory usage to ensure smooth and efficient performance, even on resource-constrained devices.
4. User-Friendly Interface: The mobile app and web front-end should offer an intuitive and user-friendly interface for users to monitor the system, access alerts, and manage settings easily.
5. Cloud Deployment: The system should be successfully deployed on a free cloud server, making it accessible and scalable for potential users.

The successful achievement of these goals will indicate that the project has developed a reliable and effective human activity detection system, particularly in dark or low-light conditions. Meeting these criteria will demonstrate that the proposed approach using the Dark Temporal Consistency Model (DTCM) neural network architecture has outperformed existing methods and delivered tangible benefits in terms of accuracy, efficiency, and user experience.

6. Preliminary Bibliography:

Github repositories:
https://github.com/sahanediriweera/SLIoT-CCTV-Human-Detection
Youtube:
https://www.youtube.com/watch?v=Ro36g2PEkBo
State of the art HAR:
https://www.v7labs.com/blog/human-activity-recognition
HAR in dark lighting:
https://xuyu0010.github.io/arid.html
http://cvpr2022.ug2challenge.org/track2.html
Papers:
DarkLight Networks for Action Recognition in the Dark