**Human Action Recognition using key point detection and MobileNetV2 Deep learning.**

**Name: Pavanteja Turaka Student ID: 21080899**

**Aim of the project**

The objective of this project is to create a resilient Human Action Recognition (HAR) system by using key point detection based on Alpha Pose model and the MobileNetV2 deep learning architecture. This system will precisely recognize and categorize human motions from images by using alpha pose to monitor and analyse human body movements and MobileNetV2 for efficient and effective extraction and classification of features. The project aims to attain a high level of precision in recognizing actions in real-time, while also ensuring computing efficiency.

**Research questions:**

* How can key point detection accuracy be optimized when integrated with MobileNetV2 for human action recognition?
* How does the integration of temporal information from key point sequences improve the action recognition capabilities of a MobileNetV2-based model?
* What are the performance trade-offs between accuracy and computational efficiency when using MobileNetV2 for key point-based human action recognition?

**Objectives of project:**

* Perform an extensive examination of the existing literature on Human Action Recognition (HAR), with a specific emphasis on identifying crucial points and exploring deep learning models, namely MobileNetV2.
* Examine current progress, emphasize areas of study that need more investigation, and provide a solid theoretical basis.
* Create a framework for key point identification by using sophisticated algorithms to accurately identify crucial moments in human motions.
* Integrate the MobileNetV2 model into the key point detection framework to provide accurate activity categorization for action identification.
* Assess the performance of MobileNetV2 in comparison to other models.
* Assess the performance of the HAR system in practical situations by conducting experiments with different datasets, using data augmentation techniques, and optimizing hyperparameters.
* Provide an elaborate report including the study methodology, experimental setting, findings, and conclusions.

**Background**

Human Action Recognition (HAR) seeks to automatically analyse and identify actions from video sequences. The need for automated human behaviour interpretation has driven significant interest in HAR for applications like video indexing, biometrics, surveillance, and security. Zhu et al. (2018) proposed a semi-supervised deep learning method using DLSTM on labelled and unlabelled data, employing smartphone inertial sensors to analyse local relationships. They achieved cutting-edge results on the UCI dataset using a DNN. Chelli and Pätzold (2019) developed a system to detect seven actions, including falls, using acceleration and angular velocity data from cell phones. Their system achieved accuracies up to 97.7% with EBT, with perfect fall detection by QSVM and EBT. Mohan et al. (2019) introduced a technique using PCA and CNN to detect abnormal behaviours in public spaces, achieving superior results on datasets like UCSD and UMN, with an 89% precision rate for identifying illicit activities via CCTV. Xia and Li (2021) utilized a pre-trained VGG-16 model with a temporal attention mechanism and LSTM to detect abnormal activities in video frames, outperforming existing methods.

**Data Management Plan**

**About dataset:**

Human Action Recognition (HAR) seeks to comprehend human behaviour and provide a categorical designation to each action. Due to its diverse variety of applications, computer vision has been garnering growing interest in the industry. The collection contains 15 distinct categories of Human Activities. The dataset consists of around 12,000 labelled photos, which includes the validation images. Each photograph is assigned to a single human activity category and is stored in a separate folder according to its labelled class.

The dataset contains images of various people doing 15 different activities, like sleeping, talking, etc. The shape and size of the image are different from each other, necessitating reshaping for the algorithm. The dataset in total has a size of 350 MB, and each image is between 10 kb and 25 kb, depending on its quality.

The data is downloaded from Kaggle and is publicly available for download and using in research projects.

Link: <https://www.kaggle.com/datasets/meetnagadia/human-action-recognition-har-dataset/data>

A person riding a bicycle

Description automatically generated A person holding a cell phone

Description automatically generated

Figure 1: Sample images.

**Document control:**

GitHub Link:

I will be using the above GitHub link for updating and version-controlling the documents. The repository will contain all the documents and code related to the project. The updating frequency will be once a week; the whole code and documentation will be updated, and the records will also be logged in Herts logbook. I will store the second copy of the code and documentation on my personal storage drive. I will also add my supervisor as a viewer to the GitHub repository, enabling him to monitor the progress of the work. Additionally, I will maintain the repository's public access for other collaborators.

The readme file will contain important information about who the code is written for, the libraries it requires for execution, and other requirements.

**Ethical requirements**

* The data complies with GDPR standards, ensuring that all personal data is handled in accordance with the EU's stringent data protection regulations.
* The project adheres to the University of Hertfordshire's ethical guidelines, guaranteeing that all research activities are conducted with integrity and respect for ethical principles.
* We have obtained permission to utilize the data for my intended research purposes, ensuring that our use of the data is authorized and appropriate.
* There is a strong assurance that the data was gathered ethically by the original collectors, ensuring that the methods of data collection, collation, and creation met ethical standards.

**Project plan**

**Ghant Chart**

Table 1: Ghant Chart

A blue and white calendar with a graph

Description automatically generated with medium confidence

The project is divided into 9 tasks, which are planned to be achieved in a span of 3 months. These tasks aim to achieve the objectives of the project and keep track of it.

* PDM submission: Initial draft due by first week of July, final document by June 14th, a critical milestone.
* Ethics quiz: Preparation and attendance scheduled over 4 to 5 days.
* Literature review: Focus on human action recognition papers, define search criteria, identify gaps in research.
* Data collection: Download datasets from provided links.
* EDA: Analyze image data, including histogram intensity analysis and dataset balance checks.
* Data preprocessing: Improve image quality based on EDA insights, balance dataset, adjust brightness, apply augmentation.
* Model development: Create and train model using preprocessed dataset, evaluate using various metrics.
* Model testing: Compare models under different conditions (e.g., without key point detection, varying weights) to select best performing model.
* Report writing: Document project findings, including methodology, results, and conclusions.

**Reference**

Xia, L. and Li, Z., 2021. A new method of abnormal behaviour detection using LSTM network with temporal attention mechanism. The Journal of Supercomputing, 77(4), pp.3223-3241.

Link: <https://link.springer.com/article/10.1007/s11227-020-03391-y>

Mohan, A., Choksi, M. and Zaveri, M.A., 2019, July. Anomaly and activity recognition using machine learning approach for video-based surveillance. In 2019 10th international conference on computing, communication and networking technologies (ICCCNT) (pp. 1-6). IEEE.

Link: <https://ieeexplore.ieee.org/abstract/document/8944396>

Chelli, A. and Pätzold, M., 2019. A machine learning approach for fall detection and daily living activity recognition. IEEE Access, 7, pp.38670-38687.

Link: <https://ieeexplore.ieee.org/abstract/document/8672567>

Zhu, Q., Chen, Z. and Soh, Y.C., 2018. A novel semi supervised deep learning method for human activity recognition. IEEE Transactions on Industrial Informatics, 15(7), pp.3821-3830.

Link: <https://ieeexplore.ieee.org/abstract/document/8586957>