

# Assignment#1 : Action Recognition Using Deep Learning Models

**Submission/viva date: 31-10-2024**

## Objective:

The goal of this assignment is to explore and implement state-of-the-art deep learning models for action recognition in video data. You will be working with a publicly available dataset to identify different human actions from video sequences using recent deep learning architecture (like we discussed in our class lectures)

---

## Tasks:

### 1. Dataset Exploration:

- Choose a publicly available dataset for action recognition. Some popular options include:
  - UCF101
  - HMDB51
  - Kinetics
- Load the dataset, explore its structure (number of classes, number of videos per class, video dimensions), and perform any necessary preprocessing (e.g., resizing, normalization).

### 2. Model Implementation:

- **Option 1: CNN + RNN (LSTM or GRU) Model**
  - Extract spatial features from individual frames using a pre-trained CNN (e.g., ResNet or VGG).
  - Use an RNN (LSTM or GRU) to capture temporal dependencies in the video.
- **Option 2: 3D CNN Model**
  - Implement or use a pre-trained 3D CNN model (e.g., C3D or I3D) to recognize actions from video sequences directly.
- Train the model using the dataset and report the accuracy, precision, recall, and F1-score for each class.

### 3. Report and Presentation:

- Submit a detailed report documenting your approach, dataset exploration, model selection, and results.
- Include the following sections:

- Introduction (problem description, motivation).
  - Dataset (overview, preprocessing steps).
  - Methodology (detailed architecture, model implementation).
  - Results (quantitative and qualitative analysis).
  - Discussion (interpretation of results, challenges, and future improvements).
  - Conclusion (summary and key takeaways).
- Prepare a short presentation (5-10 slides) summarizing your findings, including visuals of the training process, model performance, and sample predictions.

---

#### Evaluation Criteria:

- **Correctness (40%):** Proper implementation of action recognition models (CNN, RNN, 3D CNN).
- **Performance (30%):** Model accuracy, evaluation metrics, and any improvements over baseline models.
- **Creativity (20%):** Novel strategies to enhance model performance (e.g., augmentation, transfer learning).
- **Clarity of Report (10%):** Well-written report with clear explanations of methodology and results.

---

#### Submission Instructions:

- Submit your code (in Jupyter notebooks or Python scripts) and the detailed report (PDF format).
- Ensure that your code is well-commented, easy to follow, and reproducible.
- Include visualizations (e.g., graphs, confusion matrix) to support your results.
- Submit your files via the course submission portal by the due date.

---

#### Optional Bonus:

- **Real-time Action Recognition:** Implement a real-time action recognition system using a webcam or video stream. Use a lightweight model or apply pruning techniques to reduce model size and inference time.
-