**Project Stage 1**

**Report on**

### AUTOMATED HUMAN ACTION RECOGNITION AND CLASSIFICATION FRAMEWORK USING LONG RECURRENT CONVOLUTIONAL NETWORK

**Submitted in Partial fulfillment of requirements for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**In**

## COMPUTER SCIENCE AND ENGINEERING (AI & ML)

**By**

**ATCHULA HARSHANTH 20BD1A6603**

**GANDESREE SAI ABHINAV 20BD1A6614**

**GADE RANGA HRISHIKESH 20BD1A6613**

**BURRAVENI MAHARSHI 20BD1A6606**

**Under the guidance of**

**Mr. M. SRINIVAS**

***Assistant Professor, Department of CSE (AI & ML)***

******

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (AI & ML)**

|  |  |
| --- | --- |
| |  | | --- | | **KESHAV MEMORIAL INSTITUTE OF TECHNOLOGY (AN AUTONOMOUS INSTITUTION) Accredited by NBA & NAAC, Approved by AICTE, Affiliated to JNTUH.** Narayanaguda, Hyderabad, Telangana-29 **2023-24** | |

**KESHAV MEMORIAL INSTITUTE OF TECHNOLOGY   
(AN AUTONOMOUS INSTITUTION)  
Accredited by NBA & NAAC, Approved by AICTE, Affiliated to JNTUH.**

### Narayanaguda, Hyderabad, Telangana-29

# DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING (AI & ML)

**CERTIFICATE**

This is to certify that this is a bonafide record of the project report titled **“'Automated Human Action Recognition and Classification Framework Using Long Recurrent Convolutional Network”** which is being presented as Project by

* + 1. **ATCHULA HARSHANTH 20BD1A6603**
    2. **GANDESREE SAI ABHINAV 20BD1A6614**
    3. **GADE RANGA HRISHIKESH 20BD1A6613**
    4. **BURRAVENI MAHARSHI 20BD1A6606**

In partial fulfillment for the award of the degree of Bachelor of Technology in CSE (AI & ML) affiliated to the Jawaharlal Nehru Technological University Hyderabad, Hyderabad

**Internal Guide Head of Department**

**(M Srinivas) (Dr. TVG Sridevi)**

Submitted for Viva Voce Examination held on \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_



# Vision of KMIT

* + To be the fountainhead in producing highly skilled, globally competent engineers.
  + Producing quality graduates trained in the latest software technologies and related tools and striving to make India a world leader in software products and services.

# Mission of KMIT

* + To provide a learning environment that inculcates problem solving skills, professional, ethical responsibilities, lifelong learning through multi modal platforms and prepares students to become successful professionals.
  + To establish an industry institute Interaction to make students ready for the industry.
  + To provide exposure to students on the latest hardware and software tools.
  + To promote research-based projects/activities in the emerging areas of technology convergence.
  + To encourage and enable students to not merely seek jobs from the industry but also to create new enterprises.
  + To induce a spirit of nationalism which will enable the student to develop, understand India's challenges and to encourage them to develop effective solutions.
  + To support the faculty to accelerate their learning curve to deliver excellent service to students.

**VISION & MISSION OF DEPARTMENT**

**Vision of the Department**

To be among the region's premier teaching and research Computer Science and Engineering departments producing globally competent and socially responsible graduates in the most conducive academic environment.

**Mission of the Department**

* To provide faculty with state-of-the-art facilities for continuous professional development and research, both in foundational aspects and of relevance to emerging computing trends.
* To impart skills that transform students to develop technical solutions for societal needs and inculcate entrepreneurial talents.
* To inculcate an ability in students to pursue the advancement of knowledge in various specializations of Computer Science and Engineering and make them industry-ready.
* To engage in collaborative research with academia and industry and generate adequate resources for research activities for seamless transfer of knowledge resulting in sponsored projects and consultancy.
* To cultivate responsibility through sharing of knowledge and innovative computing solutions that benefits the society-at-large.
* To collaborate with academia, industry and community to set high standards in academic excellence and in fulfilling societal responsibilities.

# PROGRAM OUTCOMES (POs)

1. **Engineering Knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem Analysis:** Identify formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences
3. **Design/Development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct Investigations of Complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern Tool Usage:** Create select, and, apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The Engineer and Society:** Apply reasoning informed by contextual knowledge to societal, health, safety. Legal und cultural issues and the consequent responsibilities relevant to professional engineering practice.
7. **Environment and Sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of engineering practice.
9. **Individual and Teamwork:** Function effectively as an individual, and as a member or leader in diverse teams and in multidisciplinary settings.
10. **Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project Management and Finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-Long Learning:** Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

**PROGRAM SPECIFIC OUTCOMES (PSOs)**

**PSO1**: An ability to analyse the common business functions to design and develop

appropriate Computer Science **solutions** for social upliftment.

**PSO2**: Shall have expertise on the evolving technologies like Mobile Apps, CRM, ERP,

Big Data, etc.

**PROGRAM EDUCATIONAL OBJECTIVES (PEOS)**

**PEO1:** Graduates will have successful careers in computer related engineering fields or will be

able to successfully pursue advanced higher education degrees.

**PEO2:** Graduates will try and provide solutions to challenging problems in their profession by

applying computer engineering principles.

**PEO3:** Graduates will engage in life-long learning and professional development by rapidly

adapting changing work environment.

**PEO4:** Graduates will communicate effectively, work collaboratively and exhibit high levels

of professionalism and ethical responsibility.

# PROJECT OUTCOMES

**O1:** Accurately detect motion and control the appliances accordingly.

**O2:** Allow control of appliances remotely.

**O3:** Change the state of appliances instantly.

**O4:** Work seamlessly over the internet.

**MAPPING PROJECT OUTCOMES WITH PROGRAM OUTCOMES**



|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **PO** | **PO**  **1** | **PO**  **2** | **PO**  **3** | **PO**  **4** | **PO**  **5** | **PO**  **6** | **PO**  **7** | **PO**  **8** | **PO**  **9** | **PO10** | **PO**  **11** | **PO**  **12** |
| **O1** | H | H | H | H | M | H | L | H | H |  | M | M |
| **O2** | H | H | H | H | H | H | M | M | H |  | M | M |
| **O3** | H | H | H | H | H | M | M | H | H |  | M | M |
| **O4** | H | H | H | H | M | H | M | H | H |  | M | M |



**L – Low M –Medium H– High**

**PROJECT OUTCOMES MAPPING**

**PROGRAM SPECIFIC OUTCOMES**



|  |  |  |
| --- | --- | --- |
| **PSO** | **PSO 1** | **PSO2** |
| **P1** | H | H |
| **P2** | H | M |
| **P3** | M | M |
| **P4** | H | H |

**PROJECT OUTCOMES MAPPING**

**PROGRAM OBJECTIVES**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PEO** | **PEO1** | **PEO2** | **PEO3** | **PEO4** |
| **P1** | H | H | M | H |
| **P2** | H | H | M | H |
| **P3** | H | H | M | H |
| **P4** | H | H | M | H |

# DECLARATION

We hereby declare that the results embodied in the dissertation entitled **“**Automated Human Action Recognition and Classification Framework Using Long Recurrent Convolutional Network'' has been carried out by us together during the academic year 2023-24 as a partial fulfillment of the award of the B. Tech degree in CSE (AI & ML) from JNTUH. We have not submitted this report to any other university or organization for the award of any other degree.

### Student Name Roll No.

**ATCHULA HARSHANTH 20BD1A6603**

**GANDESREE SAI ABHINAV 20BD1A6614 GADE RANGA HRISHIKESH 20BD1A6613**

**BURRAVENI MAHARSHI 20BD1A6606**

# ACKNOWLEDGEMENT

We take this opportunity to thank all the people who have rendered their full support to our project work. We render our thanks to **Dr. B L Malleswari**, Principal who encouraged us to do the Project.

We are grateful to **Mr. Neil Gogte**, Founder & Director, **Mr. S. Nitin,** Directorfor facilitating all the amenities required for carrying out this project.

We express our sincere gratitude to **Ms. Deepa Ganu**, Director Academic for providing an excellent environment in the college.

We are also thankful to **Dr. TVG Sridevi**, Head of the Department for providing us with time to make this project a success within the given schedule.

We are also thankful to our guide **Mr. M.Srinivas**, for his valuable guidance and encouragement given to us throughout the project work.

We would like to thank the entire CSE(AI & ML) Department faculty, who helped us directly and indirectly in the completion of the project.

We sincerely thank our friends and family for their constant motivation during the project work.

## Student Name Roll No.

**ATCHULA HARSHANTH 20BD1A6603**

**GANDESREE SAI ABHINAV 20BD1A6614 GADE RANGA HRISHIKESH 20BD1A6613**

**BURRAVENI MAHARSHI 20BD1A6606**

**LIST OF DIAGRAMS**

|  |  |  |
| --- | --- | --- |
| S. No | Title of Diagram | Page no. |
| 1 | Proposed Model | 9 |
| 2 | Model Architecture | 10 |

**CONTENTS**

DESCRIPTION PAGE [CHAPTER - 1 1](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.2et92p0)

* 1. [INTRODUCTION 2-4](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.tyjcwt)
     1. Purpose of the project 2
     2. Problems In Existing Systems 2

[1.3 Problem Statement 3](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.3dy6vkm)

1.4 Objectives 3

1.5 [Scope and Limitations](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.2s8eyo1)  4

[CHAPTER – 2 5](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.17dp8vu)

* 1. [LITERATURE SURVEY 6](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.3rdcrjn)-7

[CHAPTER - 3 8](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.26in1rg)

* 1. METHODOLOGY 9-10
     1. Model Description [17](https://docs.google.com/document/d/1vNwnyGx9h1yMU6_gVJ2-YRqX4bTSdHyv/edit#heading%3Dh.lnxbz9)

**ABSTRACT**

Human action recognition, a vital domain within computer vision, involves the automatic identification and categorization of human activities from video data. This technology holds immense potential in applications ranging from security surveillance to healthcare monitoring and sports analysis. By analyzing and interpreting the movements and postures of individuals, computer vision systems can distinguish actions like walking, running, and gestures. The process typically involves techniques such as feature extraction, object tracking, and machine learning algorithms. Challenges in this field include variations in lighting, viewpoint, and occlusions. Ongoing research and advancements in deep learning, including convolutional neural networks (CNN) and recurrent neural networks (RNN), continue to improve the accuracy and robustness of human action recognition systems, making them increasingly valuable in diverse real-world applications. We are using Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for precise human action recognition. By combining CNN for spatial feature extraction and LSTM for temporal modeling, this approach achieves robust recognition of human activities in video data. The integration of spatial and temporal information enhances the model's ability to discern complex actions 

****

**CHAPTER-1**

**1.INTRODUCTION**

**1.1 Purpose of the project:**

The purpose of the project, "Automated Human Action Recognition and Classification Framework Using Long Recurrent Convolutional Network" is to create an intelligent system that automates the identification and classification of human actions in video data. Employing Long-Short Term Memory Recurrent Neural Networks (LRCN), the framework aims to capture both spatial and temporal features for nuanced action understanding.

The project seeks to contribute to AI research by implementing advanced techniques, facilitate real-world applications in surveillance and human-computer interaction, enhance human-machine interaction, and advance the state-of-the-art in technology.

By combining interdisciplinary approaches, the project strives to push the boundaries of current capabilities, creating efficient and applicable solutions for dynamic, real-world scenarios.

**1.2 Problems with Existing Systems:**

1. In the realm of Human Action Recognition, inefficient algorithms or poorly optimized models can lead to slow processing of video data, impacting real-time recognition capabilities. Performance bottlenecks may hinder the system's ability to analyze and classify actions promptly, affecting its overall effectiveness in dynamic scenarios such as surveillance or human-computer interaction.
2. Using outdated technologies in HAR systems may result in suboptimal model architectures and feature extraction methods. This could limit the system's ability to accurately recognize complex human actions, especially when newer techniques and advancements are available. Upgrading the technology stack is essential for staying competitive and ensuring that the HAR system leverages the latest research findings.
3. Scalability issues in a HAR system may arise when it fails to handle an increasing number of video streams or a growing dataset of diverse human actions. In dynamic environments, such as crowded public spaces, a lack of scalability can impede the system's ability to adapt to varying crowd sizes and diverse activities, affecting its reliability and applicability.
4. In the context of HAR applications, poor UX may manifest as inaccurate or delayed recognition of human actions. If the system's interfaces are not user-friendly or if there are delays in responding to actions, it could lead to frustration and reduced user satisfaction. For applications like gesture-based interfaces or virtual reality, a smooth and accurate user experience is crucial for the system's success and user adoption.

**1.3 Problem Statement:**

Human Action Recognition systems often struggle to provide satisfactory performance in real-world scenarios. Issues such as inadequate temporal modeling, limited generalization to diverse datasets, and challenges in real-time processing hinder the broader adoption of these systems. Addressing these problems is essential to advance the field and make action recognition technology more effective and applicable.

The project addresses these challenges by proposing an "Automated Human Action Recognition and Classification Framework Using LRCN." The problem statement emphasizes the need for an advanced solution that leverages Long-Short Term Memory Recurrent Neural Networks (LRCN) to enhance temporal modeling, ensuring accurate recognition of diverse human actions.

**1.4 Objective:**

To develop an Automated Human Action Recognition and Classification Framework using Long-Short Term Memory Recurrent Neural Networks (LRCN). This system aims to enhance temporal modeling, improve efficiency, scalability, and security, contributing to the advancement of action recognition technology for real-world applications in diverse and dynamic environments.

**1.5 Scope and Limitations:**

**1.5.1 Scope:**

* Real-world Implementation: The framework aims for practical application in diverse scenarios, ensuring its effectiveness in real-world contexts.
* Surveillance Systems: Emphasis on enhancing security surveillance by accurately recognizing and classifying human actions in video data.
* Human-Computer Interaction (HCI): Focus on improving user experiences through intelligent and responsive systems in HCI applications.
* Adaptability to Diverse Scenarios: The framework is designed to be adaptable to various contexts and industries, addressing the unique requirements of different applications.
* Dynamic Environments: The system targets dynamic environments, ensuring its effectiveness in scenarios with varying conditions and activities.

**1.5.2 Limitations:**

* Challenges recognizing rare or novel human actions: Difficulty with actions having limited representation in training data.
* Impaired recognition due to occlusions: Struggles with accurate recognition when action sequences are obscured.
* Compromised performance in low-light conditions: Suboptimal accuracy and reliability in environments with limited lighting.

****

**CHAPTER-2**

# 

# 2.LITERATURE SURVEY

A literature survey for Automated Human Action Recognition and Classification Framework Using Long Recurrent Convolutional Network.

1. **Human activity recognition (March 2015):**

Human Activity Recognition is one of the active research areas in computer vision for various contexts like security surveillance, healthcare, and human computer interaction. In this paper, a total of thirty-two recent research papers on sensing technologies used in HAR are reviewed. The review covers three areas of sensing technologies namely RGB cameras, depth sensors and wearable devices. It also discusses the pros and cons of the mentioned sensing technologies. The findings showed that RGB cameras have lower popularity when compared to depth sensors and wearable devices in HAR research. We classify these technologies into three main categories namely RGB camera, depth sensor and wearable device. Our review found that the popularity of RGB cameras in HAR research has dropped while both depth and wearable sensors are the substitutes. On the other hand, the use of Kinect sensor (depth sensor) into HAR system is promising. This could be a sign of the rise of Kinect as a popular sensing tool in HAR system.

**2)Human Action Recognition Based on Multi-Level Feature Fusion (2020):**

A deep learning-based method of multi- layer feature fusion for human motion recognition is presented. Combined with the Convolutional Autoencoder (CAE), it makes full use of the detailed features of the CNN middle layer. In addition, to reduce feature redundancy and improve recognition performance, a joint optimization module is designed. We evaluated the KTH data set and compared the most advanced methods. The experimental results have shown that the fusion network has strong robustness, and the average accuracy can reach 92.54%.

**3)Deep Learning for Human Action Recognition (2021):**

High speed processors available for computing are not sufficient for processing deep learning models as the tensors of very large size created after preprocessing datasets. Already datasets used for deep learning video processing are normally in big size so advance processing demands GPU processing with large memory. Many standard datasets are available for video analysis to validate designed model's accuracy. In recent times it is practically possible to build a good model using very high capacity and complex libraries like Keras, Theano and Torch on Python platform to make machines intelligent, the

proposed model achieved nearly 20% more accuracy by preprocessing the dataset as compared to the base model.

**4)Human Action Recognition from Various Data Modalities (2022):**

HAR is an important task that has attracted signiﬁcant research attention in the past decades, and various data modalities with different characteristics have been used for this task. In this paper, we have given a comprehensive review of HAR methods using different data modalities. Besides, multi-modality recognition methods, including fusion and co-learning methods, have been also surveyed. Benchmark datasets have also been reviewed, and some potential research directions have been discussed.

****

**CHAPTER-3**

# 

# 3.METHODOLOGY

# 

# 3.1.1 Visualize the Data with its Labels.

In the first step, we will visualize the data along with labels to get an idea about what we will be dealing with. We will be using the [UCF50 - Action Recognition Dataset](https://colab.research.google.com/corgiredirector?site=https%3A%2F%2Fwww.crcv.ucf.edu%2Fdata%2FUCF50.php), consisting of realistic videos taken from YouTube which differentiates this data set from most of the other available action recognition data sets as they are not realistic and are staged by actors.

## 3.1.2 Preprocessing the Data

Next, we will perform some preprocessing on the dataset. First, we will read the video files from the dataset and resize the frames of the videos to a fixed width and height, to reduce the computations and normalized the data to range [0-1] by dividing the pixel values with 255, which makes convergence faster while training the network.

**3.1.3 Split the Data into Train and Test Set**

## 3.1.4 Implement the LRCN Approach

In this step, we will implement the LRCN Approach by combining Convolution and LSTM layers in a single model. The CNN model can be used to extract spatial features from the frames in the video, and for this purpose, a pre-trained model can be used, that can be fine-tuned for the problem. And the LSTM model can then use the features extracted by CNN, to predict the action being performed in the video. But here, we will implement another approach known as the Long-term Recurrent Convolutional Network (LRCN), which combines CNN and LSTM layers in a single model. The Convolutional layers are used for spatial feature extraction from the frames, and the extracted spatial features are fed to LSTM layer(s) at each time-steps for temporal sequence modeling. This way the network learns spatiotemporal features directly in an end-to-end training, resulting in a robust model.

# 

# 

Fig 3.1 Proposed Model

We will also use [TimeDistributed](https://colab.research.google.com/corgiredirector?site=https%3A%2F%2Fkeras.io%2Fapi%2Flayers%2Frecurrent_layers%2Ftime_distributed%2F) wrapper layer, which allows applying the same layer to every frame of the video independently. So it makes a layer (around which it is wrapped) capable of taking input of shape (no\_of\_frames, width, height, num\_of\_channels) if originally the layer's input shape was (width, height, num\_of\_channels) which is very beneficial as it allows to input the whole video into the model in a single shot.

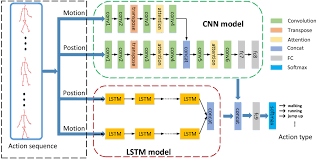


Fig 3.2 Model Architecture

## 3.1.5 Upload a video in test\_videos folder.

Run the predict\_single\_action cell.