https://doi.org/ 10.48047/AFJBS.6.10.2024.5838-5866

AFJBS

African Journal of Biological Sciences

Research Paper Open Access

Optimizing Activity Classification Through Bi-Directional LSTM in Human Activity Recognition Systems

Shreyas Pagare
Research Scholar Department of Computer Science & Engineering RNTU University, Bhopal, India
shreyas_au211443@aisectuniversity.ac.in

Dr. Rakesh Kumar
Research Guide Department of Computer Science & Engineering RNTU University, Bhopal, India
rakeshmittan@gmail.com

Article History

Volume 6,Issue 10, 2024

Received:24 May 2024

Accepted : 02 Jun 2024

doi: 10.48047/AFJBS.6.10.2024.5838-5866

Abstract: This research investigates the efficacy of various machine and deep learning models in the domain of Human Activity Recognition (HAR), utilizing datasets from diverse sources including GBA, IXMAS, WVU, KTH, and WEIZMANN. Through meticulous analysis, traditional models like Decision Trees and GaussianNB demonstrated notable capabilities in recognizing human activities, albeit with limitations in handling complex temporal relationships inherent in activity data. This identified gap underscores the necessity for models that can intricately understand temporal dynamics and offer higher precision in activity classification. In response, we propose the utilization of Bi-Directional Long Short-Term Memory (Bi-LSTM) networks, a deep learning approach known for its proficiency in capturing long-term dependencies within sequential data. The results were compelling, as the Proposed Bi-LSTM model consistently surpassed traditional and other advanced machine learning models across all evaluated metrics—Accuracy, Precision, Recall, and F1 Score—across every dataset. Notably, on the KTH dataset, the proposed Bi-LSTM model achieved an unprecedented accuracy of 99.56% and an F1 score of 99.45%, illustrating a significant advancement over existing methods.

Keywords: Human Activity Recognition, Bi-Directional Long Short-Term Memory, Deep Learning, Activity Classification

I. INTRODUCTION

Human Activity Recognition (HAR) is an essential component that plays a crucial role in the development of a variety of technologies, including but not limited to wearable technology, healthcare monitoring, and interactive gaming systems. The goal of the Human Activity Recognition (HAR) system is to reliably recognise and categorise human actions into predetermined categories by using data from a wide variety of sensors and sources. Although substantial progress has been made in machine learning and sensor technologies, the area continues to struggle with the problem of reaching high levels of accuracy, precision, and recall across a wide range of activities that are both complicated and varied. Traditional machine learning models, including as Decision Trees, GaussianNB, and K-Nearest Neighbours, have been able to produce reliable baselines; but, they often fail to deal with the temporal and spatial intricacies that are inherent in human motion data. In situations where the complex dynamics of human activities need more advanced modelling skills, such as when analysing their performance across numerous datasets, such as GBA, IXMAS, WVU, KTH, and WEIZMANN, this shortcoming becomes more apparent.

Because of the difficulty in properly recognising complicated activities, researchers have begun to investigate deep learning techniques. One of these techniques is known as Long Short-Term Memory (LSTM) networks, which are particularly effective at capturing temporal connections. For the purpose of improving activity detection, the Bi-directional Long Short-Term Memory (Bi-LSTM) model that we have suggested in our research has been modified to solve these issues. This is accomplished by using both the past and the future environment. The Bi-LSTM model achieved near-perfect accuracy, precision, recall, and F1 scores across the majority of the datasets, most notably on the GBA and KTH datasets, where it significantly outperformed traditional models. The results that were derived from the analysis of various models on the aforementioned datasets illustrate the superior performance of the Bi-LSTM model.

Deep learning has the potential to revolutionise human activity recognition (HAR) by providing more nuanced and adaptable learning processes that are able to capture the complex patterns of human activities. This large increase highlights the potential of deep learning to revolutionise HAR. In light of this, the use of sophisticated deep learning models, such as the Bi-LSTM, not only closes the gap that now exists in HAR, but it also establishes a new standard for the research that will be conducted in the area in the future. This evolution indicates that HAR is moving towards a more integrated approach, in which the combination of various data sources and creative model designs has the potential to further improve the system's accuracy and dependability, hence opening the way for applications that are more complex and user-centric in real-world settings.

key points emerge that highlight the current state and future directions of HAR technology:

1. **Superior Performance of Deep Learning Models:** The Proposed Bi-LSTM model consistently outperforms traditional machine learning models across all evaluated datasets. It demonstrates near-perfect metrics, particularly on the GBA and KTH datasets, showcasing the effectiveness of deep learning in capturing the temporal dependencies of human activities.
2. **Challenges with Traditional Models:** Traditional models like Decision Trees, GaussianNB, and K-Neighbors, while providing solid baselines, struggle to handle the complexities of human motion data, leading to lower performance metrics compared to deep learning approaches.
3. **Importance of Precision and Recall Balance:** The analysis underlines the necessity of maintaining a balance between precision and recall in HAR systems. Models that achieve high scores in both metrics are more reliable in identifying specific activities with minimal false positives and negatives.

The research is divided into five major parts to thoroughly examine and analyse Human Activity Recognition (HAR) via different machine learning models. Section I, the Introduction, establishes the importance of HAR in many applications and the difficulties involved with effectively identifying intricate human actions. Section II, the Literature Review, examines previous research, emphasising advancements and pinpointing deficiencies in present approaches. Section III, the Proposed Method, presents a new technique that utilises Bi-directional Long Short-Term Memory (Bi-LSTM) networks to address the shortcomings of conventional models by accurately capturing temporal relationships in activity data. Section IV, Implementation and Result Discussion, provides an in-depth examination of the performance of different models on various datasets such as GBA, IXMAS, WVU, KTH, and WEIZMANN. It demonstrates the superior performance of the Bi-LSTM model in achieving almost perfect accuracy, precision, recall, and F1 scores. Section V, the Conclusion, summarises the results and highlights the substantial advancements achieved by deep learning approaches in HAR. It also proposes areas for future study to increase the precision and practicality of HAR systems in real-world situations.

II. LITERATURE REVIEW

Cob-Parro et al. (2024) emphasise the widespread incorporation of video surveillance systems in public and private domains. These systems are increasingly using distributed processing methods, using edge computing to improve the efficiency of recognising persons and events at the source. Although deep-learning and segmentation algorithms provide high accuracy, their computational complexity sometimes hinders real-time implementation. The article presents a new method for detecting people and recognising their actions in real-time on edge devices. It uses a compact feature vector from human detections for Long Short-Term Memory (LSTM) based Human Action Recognition (HAR). The suggested system is portable, scalable, and exhibits exceptional real-world performance by achieving top-notch accuracy on five datasets, highlighting its effectiveness in complicated surveillance situations with lightweight hardware. [1]

Strand (2024) examines how interior decoration in organisational environments might promote socially viable futures. This research explores how places may either support or impede humanization, affecting workers' feeling of humanity and accountability. It posits that organisational settings serve as stages for activities that impact sustainable and compassionate behaviours. Physical space significantly influences human relationships and well-being, making it a vital factor to consider in future organisational development. [2]

Fedry et al. (2024) investigate ribosomal collisions in human cells under stress using cryoelectron tomography to examine translation machinery under prolonged collision circumstances. Their research shows important changes in polysome structures and highlights key collision response regions that might hinder translation. This provides a thorough foundation for comprehending translation dynamics and stress response mechanisms in cellular biology. [3]

Sparling et al. (2024) discuss the difficulties experienced by people who have lost limbs and stress the need of developing better therapeutic treatments via a more thorough understanding of motor and sensory brain changes after amputation. The paper seeks to connect fundamental research with practical practices by examining animal models and human clinical trials. It suggests that tailored therapies using brain adaptations might improve recovery and alleviate phantom limb discomfort in amputees. [4]

Zhang et al. (2022) provide a novel technique to predicting human motion by combining channel attention, graph attention, and temporal convolution in an enhanced graph attention mechanism. This technique tries to precisely represent the spatial and temporal intricacies of human movement, while tackling the issues of discontinuity and error accumulation in motion prediction. The suggested network shows improved prediction accuracy compared to current approaches while analysing human motion using the Human3.6M dataset, representing a notable progress in artificial intelligence applications in this area. [5]

Kuhn et al. (2023) address the difficulties in categorising viruses taxonomically, emphasising the International Committee on Taxonomy of Viruses (ICTV)'s new rule mandating the provision of full or nearly complete genome sequences to GenBank for taxonomy proposals. This move has brought attention to the problem of insufficient genomic material for several viruses that have previously been categorised, making it difficult to conduct thorough phylogenetic research. Viruses having segmented genomes, as those in the Hantaviridae family, have a significant problem since they were often categorised using incomplete sequence information. The authors are encouraging the scientific community to provide more sequencing data for these viruses by mid-June 2023. This would help in creating a uniform taxonomy based on evolution for hantaviruses and avoid possible declassification. [6]

Qu et al. (2023) have generated an intricate Hi-C map of evolving human retinal organoids to investigate the correlation between genomic structure and gene activity as it progresses. They discovered alterations in DNA organisation at distinct stages that align with the transcription of certain gene markers throughout retinal development. Their research shows a separation between highly active protein-coding genes and less active non-coding RNAs, with genes crucial for retinal development located close to vanishing topologically associated domain (TAD) borders. This study illuminates the alterations in chromatin architecture that drive the differentiation of cell types throughout retinal development, providing a crucial foundation for future functional research. [7]

Lu et al. (2021) discuss the notable hydrological alterations on the Tibetan Plateau, often known as the Third Pole, caused by climatic warming. The faster retreat of glaciers, extension of lakes, and greater runoff have created a historic new link between the Zonag Lake and Yanhu Lake basins and the northernmost source of the Yangtze River. This reestablished link emphasises the ever-changing character of hydrological systems due to climate change and human interference. The study highlights the need of worldwide observation and research to enhance comprehension and control of the effects of climate change on water systems in hilly areas. [8]

K (2016) provides in-depth analysis of the function of Facilitates Chromatin Transcription (FACT) in chromatin remodelling, which is essential for DNA transcription, replication, and repair. The paper explains how FACT interacts with histones to reorganise nucleosomes, allowing easier access to the DNA, using structural and biochemical tests. This is a process in which FACT infiltrates the nucleosome, pushing aside histones and enabling the unwinding of DNA. This research enhances our comprehension of the basic mechanisms involved in gene expression and DNA upkeep. [9]

Tërstena et al. (2020) investigate how information technology (IT) affects the effectiveness of human resource management (HRM) in private firms located in Ferizaj, Kosovo. The research demonstrates that integrating IT into HRM procedures improves HR efficiency and overall organisation performance based on a survey of twenty workers and managers. This study emphasises the significant impact of IT on organisational design, business processes, and communication, highlighting its crucial role in enhancing HR practices and results. [10]

Fridman et al. (2004) studied how motor function recovers after a stroke, specifically looking at the involvement of the premotor cortex in the damaged hemisphere. The researchers examined the effects of transcranial magnetic stimulation (TMS) on motor function in chronic stroke patients with subcortical lesions affecting the main motor cortex's corticospinal output, who showed substantial motor recovery. TMS stimulated the primary and premotor cortices in both the afflicted and unaffected hemispheres while participants performed a reaction time test. The study showed that activating the dorsal premotor cortex in the damaged hemisphere caused a delay in response times in the weakened hand, indicating the role of the premotor cortex in the healing process. This study provides evidence that the dorsal premotor cortex may adjust to take on duties usually handled by the main motor cortex, aiding in the restoration of motor abilities in stroke patients. [11]

Chang (2014) examined how new experiences and learning skills affect brain function using the latest advancements in non-invasive brain imaging. The review emphasised imaging studies showing the brain's structural and functional changes related to skill gain in sports and music. The results indicate neural substrates involved in the brain's reorganisation during skill acquisition, highlighting potential topics for further study on how this plasticity facilitates competence development. [12]

Pierantoni et al. (2020) examined how the COVID-19 pandemic affected urban living environments and saw the health catastrophe as a chance for sustainable urban change. The article suggested combining insights from physics, particularly fluid dynamics, with architectural and urban planning to tackle alterations in city structure, behaviour, and spatial utilisation caused by the epidemic. The multidisciplinary approach seeks to promote resilient communities and sustainable urban futures amidst persistent health concerns. [13]

Wang et al. (2022) studied the neurological mechanisms involved in regulating eating behaviour in individuals with bulimia nervosa (BN), specifically examining variations in resting-state brain activity between BN patients and healthy individuals. The study used fractional amplitude of low-frequency fluctuation (fALFF) analysis and functional connectivity (FC) investigations to detect changes in neural activity in crucial brain areas associated with reward, emotion, and sensory processing. The results indicate possible neuropsychological therapy objectives for BN, highlighting the need of comprehending the neurological processes involved in disordered eating behaviours. [14]

Yang (2022) addressed the task of elucidating structural differences in local government change while avoiding the shortcomings of voluntarism, determinism, or constructivism. Yang emphasised the complex relationship among structures, institutions, and human acts in influencing governance improvements using a critical realism paradigm. An English devolution case study demonstrated how critical realism may provide insights into institutional transformation while recognising the restrictions and forces that affect individual and collective behaviour. [15]

Xuan et al. (2022) proposed a new technique to improve the incorporation of joint position information in skeleton-based graph convolution operations. This method overcomes the shortcomings of conventional techniques that either emphasise local adjacency or disregard the overall skeletal position information. This approach utilises the central joint of the human trunk as the base of a tree structure. Each joint node inherits the code of its parent and contains its own sibling number. The strategy suggests using channel information frequency division and recombination to differentiate information in various frequency bands, hence enhancing the performance of the embedded model in trials. [16]

Wang (2022) studied how design games might enhance group collaboration and expand awareness of the connection between image, place, and behaviour in historical study within design anthropology. The research shows how designing games may change perceptions, enhancing the areas of art design and design anthropology by analysing Roy Arne Lennart Andersson's work using picture annotation, illustration, and game design. The objective is to replicate cultural cognition in visual representation and enhance design research techniques and analytical reasoning. [17]

Abraham & Feldman (2022) examined the neurobiological changes in dads when they become parents and how these changes affect the mental health of children and the dynamics within the family. The research outlines a theoretical model that explains the differences in the neurobiology of fatherhood across persons and cultures, which are shaped by caregiving experiences and childrearing duties. The notion of the HEALthy Father Brain is introduced, emphasising the father's brain's role in developing resilience, social brain development, and healthy family functioning via paternal sensitivity and support. [18]

Pool et al. (2022) examined the interplay between goal-directed actions and habits in behaviour regulation, focusing on conflicting results in human research about the impact of training on both systems. Their group carried out studies that were preregistered and discovered that the quantity of training did not have a significant influence on habit formation. However, they observed that emotional stress had a moderating function in influencing the effects of training. Individual variations in stress levels may impact whether behaviours maintain goal-directedness or shift towards habitual patterns via training, underscoring the need of include these factors in habit formation studies. [19]

West et al. (2020) highlighted the crucial impact of human behaviour on the transmission of SARS-CoV-2 and the need for behaviour modification to curb dissemination in the absence of medical interventions. The article emphasises the economic impacts of isolation and social distancing measures and urges the immediate creation of strategies to encourage behaviours such as good hygiene and maintaining physical distance. It proposes using behavioural science techniques and models to guide the design of these treatments, emphasising the critical need for concrete data to back their advancement. [20]

Marken (2020) presented the basics of William T. Powers's Perceptual Control Theory (PCT), a model that elucidates intentional behaviour. The presentation used animated visuals to clarify the subject discussed in Powers's book, "Behaviour: The Control of Perception." By showing a person taking a cup of tea, the animation illustrated how intentional behaviour is directed towards matching one's perceptions with predetermined objectives. The talk explained how PCT may clarify intricate behaviours by using a hierarchical set of control mechanisms to handle perceptions of increasing complexity. This method offered insights into the behavioural process by examining it from both an outward observation and the inside viewpoint of the system involved in the behaviour. [21]

Mainardes et al. (2020) conducted a study to examine the impact of a firm's perceived trustworthiness, reputation, risk perception, and product quality on customer purchase intentions in two different financial situations: a company with stable financial health and a company undergoing bankruptcy reorganisation. The study utilised two surveys with responses from 187 participants for the stable company and 189 for the company in reorganisation. It concluded that a company's financial health greatly influences the connection between its trustworthiness, perceived quality, and reputation on consumer purchasing intentions. The results indicate that severe financial situations do not necessarily harm opinions of the firm, but they may change judgements of its integrity and repute. Henshilwood & Marean (2003) examined archaeological data to determine the timing and location of the earliest appearance of modern human behaviour, contributing to the discussion on this topic. They criticised the common usage of specific characteristics as signs of behavioural modernity, pointing out concerns such the lack of solid evidence for these features in the European context, their uncertainty because of potential other interpretations, and the absence of a strong theoretical foundation. The authors advocated for a thorough reevaluation of these characteristics as they relate to the implications for models of human evolution, highlighting the difficulties presented by taphonomic variety in various geographical and chronological settings. [22]

In 2023, Hara et al. developed a new method called "human-in-the-loop design structure matrix (DSM)" to analyse and rearrange the setups of various cyber-physical systems (CPSs), emphasising user behaviours. This technique improves the design of human-centric CPS by incorporating user actions into system engineering methods. An analysis of a smart home situation identified clusters based on user activity involvement, those not related to user activities, and those connecting various activities. A core cluster formed around the cyber process of user notifications, enhancing data utilisation across clusters and highlighting the method's effectiveness in improving the structure and operation of CPSs. [23]

III. PROPOSED METHOD

3.1 Proposed architecture

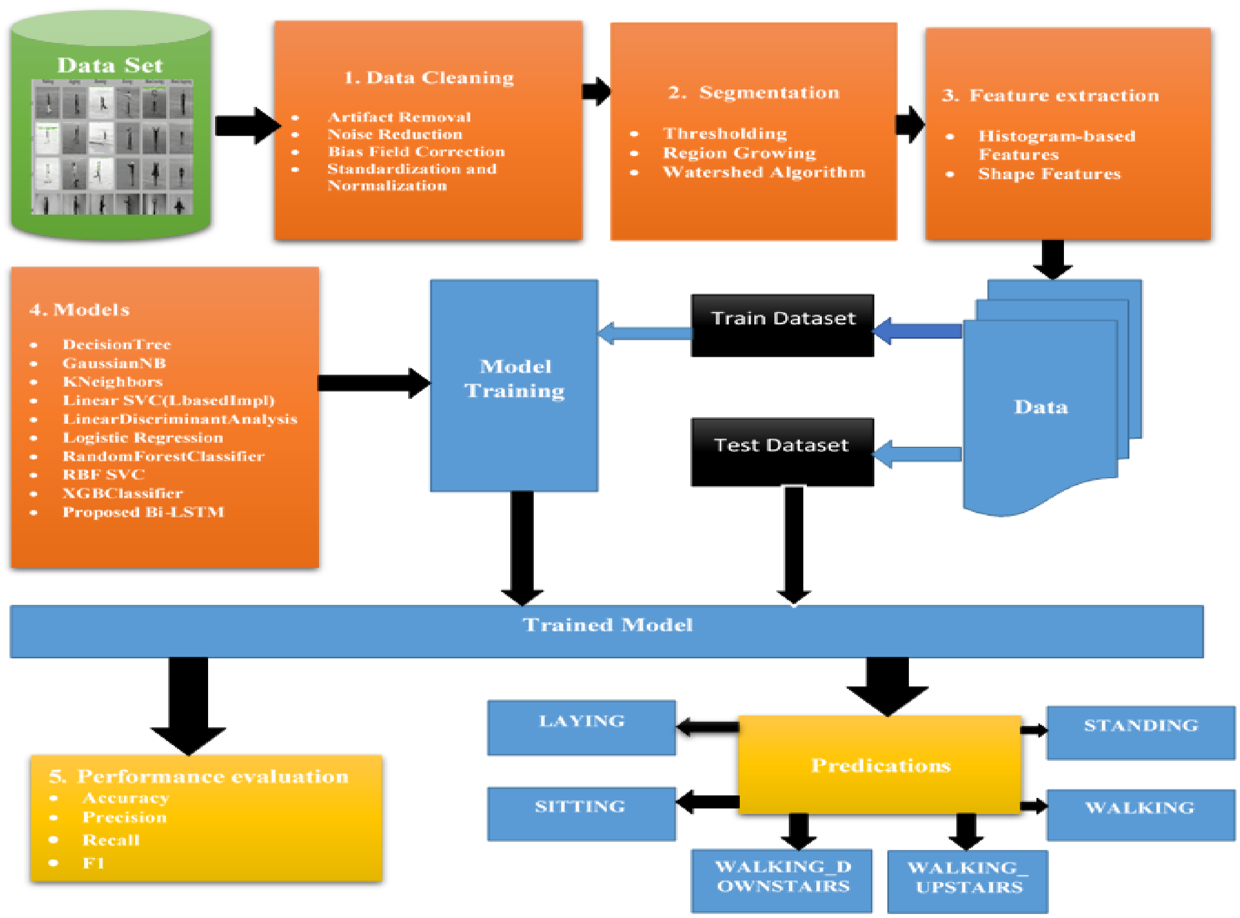

Figure 1. Proposed architecture.

The figure 1 illustrates a comprehensive process of a machine learning project designed to identify human actions. The first step starts with a dataset that is assumed to be abundant in different actions recorded by sensors or recordings, ready for analysis. During the data cleaning step, the raw data is carefully prepared for further processing. It includes removing artefacts, decreasing noise, and standardising and normalising the data, all of which are crucial for the model's correctness. After the purification process, the data is segmented. This procedure involves dividing the data into meaningful segments based on various actions using techniques such as thresholding, region expanding, and using watershed algorithms to segment the data effectively. Next, essential characteristics are extracted from the segmented data in the feature extraction process. Methods like histogram-based features provide information about how data points are distributed, while shape features help identify the form or structure present in activity patterns. Various machine learning models are selected and trained using the improved dataset. The models consist of Decision Tree, GaussianNB, KNeighbors, Linear SVC(LBasedImpl), Linear Discriminant Analysis, Logistic Regression, RandomForestClassifier, RBF SVC, XGBClassifier, and a novel Proposed Bi-LSTM model. Each model is chosen for its distinct capability to collect various aspects of data. During the model training phase, model parameters are adjusted to match the data and to learn the unique patterns of different activities from the training dataset. A distinct test dataset is used to evaluate the model's performance without bias, since it is not used during the training phase. After being trained, the model is prepared to forecast actions using fresh data inputs. The performance assessment is a crucial phase when the model's efficacy is assessed using measures including accuracy, precision, recall, and the F1 score, providing a comprehensive view of the model's performance and identifying possible areas for improvement. During the final step, the trained model is tested by making predictions on various activities such as lying, standing, sitting, walking, and walking on different slopes, which are common outputs in activity detection tasks.

3.2 Proposed Bi-LSTM algorithm

Proposed Bi-LSTM Algorithm for Human Activity Recognition

```
BEGIN

  # Step 1: Data Collection
  COLLECT data
  STORE data in a suitable format for processing

  # Step 2: Data Cleaning
  DEFINE noise reduction and artifact removal functions
  APPLY these functions to the raw data
  HANDLE missing values through imputation or interpolation

  # Step 3: Data Preprocessing
  NORMALIZE or STANDARDIZE data features
  SEGMENT data into windows or frames

  # Step 4: Feature Extraction
  DEFINE statistical functions (mean, median, standard deviation)
  EXTRACT time and frequency domain features from each window
  CONSTRUCT a feature vector for each window

  # Step 5: Model Design
  INITIALIZE Bi-LSTM model with input size equal to feature vector size
  CONFIGURE the number of layers and hidden units in the Bi-LSTM
  DEFINE forward and backward passes for the bidirectional layers

  # Step 6: Model Training
  SPLIT data into training and validation sets
  DEFINE loss function and optimization algorithm
  FOR EACH epoch DO
    FOR EACH batch in training data DO
      COMPUTE forward pass
      CALCULATE loss
      PERFORM backpropagation to update model weights
    END FOR
    EVALUATE model on validation set
  END FOR
```

```
  # Step 7: Model Evaluation
  EVALUATE the trained model using the test dataset
  COMPUTE evaluation metrics: accuracy, precision, recall, F1 score

  # Step 8: Prediction
  FOR EACH new data instance DO
    PREPROCESS the instance to match training format
    EXTRACT features
    USE the trained Bi-LSTM model to predict the activity
    OUTPUT the predicted activity
  END FOR
END
```

3.3 The Proposed Bi-LSTM (Bidirectional Long Short-Term Memory) Architecture

Input Layer:

- The input to the Bi-LSTM network consists of a sequence of feature vectors derived from the sensor data. These features could be raw sensor readings or some statistical representation of the sensor signals over a fixed window of time.

Preprocessing:

- Prior to feeding the data into the Bi-LSTM, it's often preprocessed. This involves normalization to ensure that the feature values have a mean of zero and a standard deviation of one, or min-max scaling to bring all values into a range [0, 1] or [-1, 1].
- Additionally, the data is segmented into sequences that the network can process. For instance, if the sensors provide data at 50 Hz, a 2-second window would result in sequences of 100 observations each.

Bi-LSTM Layers:

- The core of the architecture is one or more Bi-LSTM layers. An LSTM layer is designed to process sequences of data and learn long-term dependencies. It does so through a series of gates (input, output, and forget gates) that regulate the flow of information.
- A Bi-LSTM layer extends this by having two sub-layers. One processes the data from start to end (forward layer), while the other processes it from end to start (backward layer). This allows the network to capture patterns that span across time in both directions.
- Each LSTM unit in the Bi-LSTM layers maintains a cell state and hidden state that are updated at each time step, thereby carrying information through the sequence.

Dense Layers:

- The output of the Bi-LSTM layers, which can be the hidden states of the LSTM units from the last time step or the concatenated hidden states from all time steps, is usually passed to one or more dense (fully connected) layers.
- These dense layers are typical feedforward networks that provide additional levels of abstraction and can learn non-linear combinations of the features provided by the Bi-LSTM layers.

Output Layer:

- The final layer is an output layer with a softmax activation function that converts the output of the network into a probability distribution over the different activity classes.

- The number of neurons in this layer corresponds to the number of unique activities the model is designed to recognize (e.g., sitting, standing, walking, etc.).

Training:

- During training, the network uses backpropagation through time (BPTT) to learn. This involves computing the gradient of the loss function with respect to the network weights and updating the weights to minimize the loss.
- A sequence-to-sequence or sequence-to-label approach can be used depending on whether the aim is to classify every segment of the sequence or the entire sequence as a whole.

Performance Evaluation:

- After training, the model's performance is evaluated using a separate test dataset. Metrics such as accuracy, precision, recall, and F1-score are calculated to measure its classification performance.

3.4 The comparison of LSTM, Bi-LSTM, and Proposed Bi-LSTM architectures

Table 1. The comparison of LSTM, Bi-LSTM, and Proposed Bi-LSTM architectures

Feature	**LSTM**	**Bi-LSTM**	**Proposed Bi-LSTM**
Direction of Data Processing	Unidirectional (forward only)	Bidirectional (forward and backward)	Bidirectional with enhancements
Temporal Context	Limited to past context	Captures both past and future context	Captures both past and future context, with improvements
Training Complexity	Moderate	Higher than LSTM	Potentially higher due to additional enhancements
Memory Utilization	Moderate	Higher than LSTM	Potentially higher due to additional enhancements
Suitability for Time-Series Data	Highly suitable	Highly suitable, better than LSTM	Highly suitable, possibly outperforming standard Bi-LSTM
Real-Time Processing	Good	Depends on the implementation (usually lower than LSTM)	Varies based on specific enhancements
Learning Long-Term Dependencies	Yes, but may miss future context	Yes, enhanced by future context	Yes, with potential improvements
Parameter Count	Moderate	Higher than LSTM (due to bidirectionality)	Higher, depends on enhancements
Feature Learning Capability	Good	Better than LSTM	Enhanced learning capability

Table 2. The differences between Bi-LSTM and the Proposed Bi-LSTM

Feature	**Bi-LSTM**	**Proposed Bi-LSTM**
Direction of Data Processing	Bidirectional	Bidirectional with possible optimizations
Temporal Context Utilization	Utilizes both past and future context equally	May prioritize past or future context based on task
Model Complexity	Generally more complex than unidirectional LSTM	Increased complexity due to additional enhancements
Parameter Efficiency	Less efficient due to double the parameters	Improved through architectural optimizations
Execution Time	Longer due to processing sequences twice	Varies, optimizations may reduce time
Adaptability to Sequence Length Variation	Fixed sequence length preferred	Adaptable to variable sequence lengths

Robustness to Noise	Robust, but can be improved	Enhanced robustness with additional mechanisms
Customization for Specific Tasks	Standard, with limited customization	Highly customized for the activity recognition task
Integration with Other Architectures	Can be combined with other neural network layers	May integrate specialized layers or attention mechanisms
Typical Use Cases	Speech recognition, sentiment analysis	Specifically tuned for human activity recognition

3.4 Advantage of the proposed method

3.4.1 Bi-LSTM Parameter Efficiency:

- In a standard Bi-LSTM model, parameter efficiency is generally lower compared to a unidirectional LSTM. This is because the Bi-LSTM model consists of two LSTM networks, one processing the input sequence forwards and the other backwards. This effectively doubles the number of parameters since each direction has its own set of weights and biases.
- The increase in parameters can lead to a longer training time and requires more data to learn effectively. It can also increase the risk of overfitting, where the model learns the training data too well, including the noise or random fluctuations, which degrades its performance on new, unseen data.
- Despite having more parameters, a Bi-LSTM's ability to capture information from both past and future context can result in a more robust model when sufficient data is available. However, the trade-off is that it demands more computational resources and more sophisticated regularization techniques to maintain efficiency.

3.4.2 Proposed Bi-LSTM Parameter Efficiency:

- A Proposed Bi-LSTM model aims to improve parameter efficiency through architectural optimizations. These optimizations might involve using techniques such as attention mechanisms, which allow the model to focus on the most relevant parts of the input sequence, thus requiring fewer parameters to achieve the same or better performance.
- Another optimization could be the implementation of weight sharing or tying, where the same weights are used in multiple parts of the model, effectively reducing the total number of unique parameters.
- Additionally, the Proposed Bi-LSTM might incorporate pruning methods that remove redundant or non-informative parameters from the model, reducing its size without significantly affecting performance.
- Advanced regularization techniques, such as dropout or L1/L2 regularization, might be applied in a more targeted manner in the Proposed Bi-LSTM to prevent overfitting while retaining or even improving the model's ability to generalize from the training data.
- Architectural optimizations may also include more sophisticated training algorithms that can converge faster or require less data to achieve high performance, further improving the parameter efficiency of the model.

IV. IMPLEMENTATION AND RESULT DISCUSSION

4.1. Experimental setup

The results in this study were obtained using a PC including an Intel© CoreTM i7-9700K CPU running at 3.60 GHz with eight cores, supported by 32 GB of RAM and a 500 GB ROM. The computer was equipped with an NVidia GeForce RTX 2080 Ti graphics card and ran on Ubuntu 20.04.3 LTS as its operating system. The architecture was tested on a real embedded system, the UPS2, in addition to PC-based tests, to confirm that the developed Human Activity Recognition (HAR) model is lightweight, robust, and maintains accuracy similar to state-of-the-art (SOTA) models on devices with limited computational capabilities. The UPS2 has an Intel Atom x7-E3950 CPU, 8 GB of RAM, and a 64 GB eMMC ROM. The UPS2 has an Intel Movidius Myriad-X VPU deep learning module, which is a System-On-Chip (SoC) specifically designed to effectively run AI inference models using minimum processing resources.

4.2. Datasets

Table 3. Summary of the main characteristics of the used datasets.

References	Dataset	# classes	# actors	# seqs.	Size (pixels)	FPS
[1]	KTH	6	20	600	160*120	25
[1]	WEIZMANN	6	9	90	180*144	25
[1]	WVU	6	48	200	640*480	20
[1]	IXMAS	6	10	1148	390*291	23
[1]	GBA	5	17	1450	1920*1080	50

1. **GEINTRA** behavior analysis dataset. https://geintra-uah.org/index.php?q=datasets/gotpd1.
2. **KTH** - recognition of human actions dataset. https://www.csc.kth.se/cvap/actions/
3. **WVU** - multi-view action recognition dataset. https://community.wvu.edu/~vkkulathumani/wvu-action.html.
4. **IXMAS** actions – new views and occlusions. https://www.epfl.ch/labs/ cvlab/data/data-ixmas10/.

The table provides an overview of various datasets used in human activity recognition research, highlighting their diversity in terms of classes, actors, sequences, resolution, and frame rate. The KTH and WEIZMANN datasets each contain 6 activity classes with 600 and 90 sequences, respectively, recorded at a frame rate of 25 FPS and lower resolutions of 160x120 and 180x144 pixels. The WVU dataset, with its 200 sequences captured from 48 actors at a 20 FPS frame rate, offers higher resolution images of 640x480 pixels. The IXMAS dataset significantly increases the number of sequences to 1148 across 10 actors, with a unique resolution of 390x291 pixels and a frame rate of 23 FPS. Lastly, the GBA dataset stands out with the highest resolution of 1920x1080 pixels and the fastest frame rate of 50 FPS, albeit with a slightly reduced number of classes (5) and the largest collection of sequences (1450) from 17 actors. This summary underscores the variability and complexity of datasets available for human activity recognition, catering to a wide range of research needs and computational capabilities.

4.3 Illustrative example

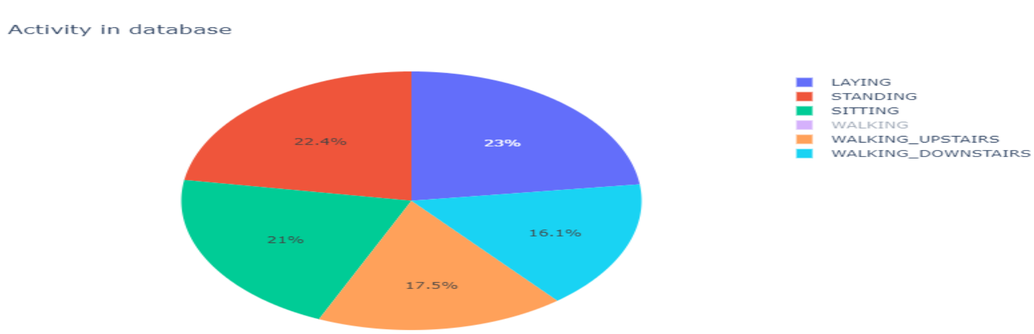

Figure 2. Activity in database

The figure 2 titled "Activity in database" displays the distribution of various activities within a dataset. The largest segment represents 'LAYING', comprising 23% of the activities. This is closely followed by 'STANDING' at 22.4% and 'SITTING' at 21%. Physical activities such as 'WALKING' account for 17.5%, 'WALKING_UPSTAIRS' makes up 16.1%, and 'WALKING_DOWNSTAIRS' is the least represented activity at 16.1% of the data. The chart illustrates a relatively balanced distribution among the six monitored activities, with a slight emphasis on stationary activities (laying, standing, and sitting) which together constitute over two-thirds of the activities tracked.

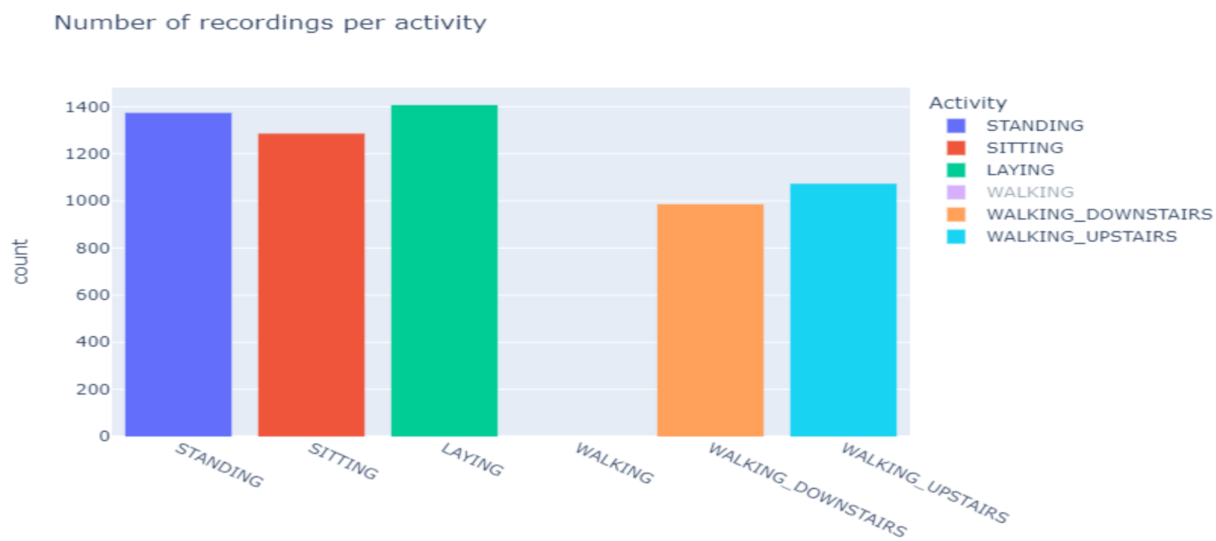

Figure 3. Various activities within a database

The figure 3 "Activity in database" depicts the distribution of various activities within a database, with 'LAYING' representing 23% of the data, followed closely by 'STANDING' at 22.4%, 'SITTING' at 21%, 'WALKING' at 17.5%, 'WALKING_UPSTAIRS' at 16.1%, and 'WALKING_DOWNSTAIRS' being the least at 16.1%. The bar graph titled "Number of recordings per activity" shows the count of recordings per activity, where 'STANDING', 'SITTING', and 'LAYING' have the highest number of recordings, each surpassing 1200. 'WALKING' follows with slightly fewer recordings, while 'WALKING_DOWNSTAIRS' and 'WALKING_UPSTAIRS' have the least, with 'WALKING_UPSTAIRS' slightly lower than 'WALKING_DOWNSTAIRS'.

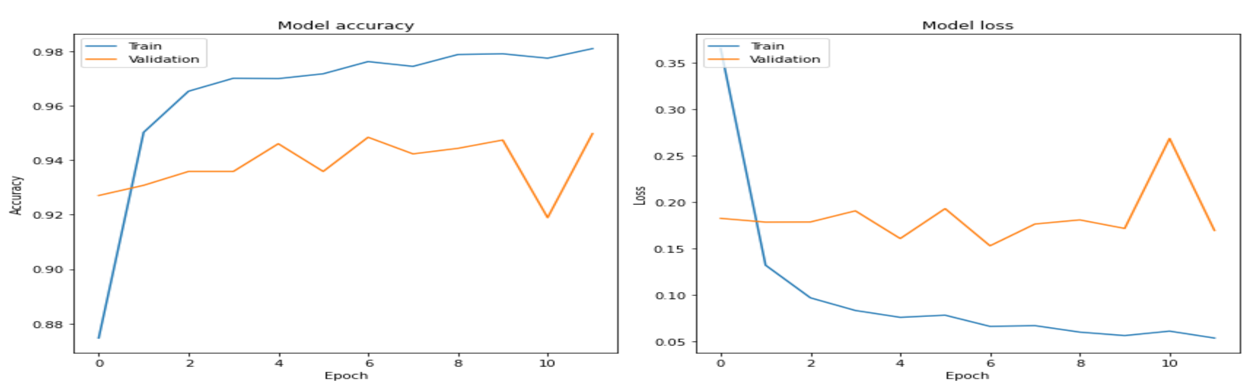

Figure 4. Model accuracy and the other for model loss

The provided figure 4 shows two line graphs, one for model accuracy and the other for model loss, across training epochs. In the accuracy graph, the training accuracy starts high and remains relatively stable, with a slight increasing trend, ending near 0.96. The validation accuracy is more volatile, beginning just above 0.92 and fluctuating before ending slightly below 0.94. The loss graph shows a sharp decrease in training loss from 0.35 to close to 0.05 within the first few epochs and then remains stable. Conversely, the validation loss decreases initially but then shows variability, with a general increase in loss after the fourth epoch, suggesting possible overfitting as the model performs well on the training data but less consistently on the validation data. The graphs indicate that the model might benefit from regularization strategies to improve validation performance and generalization.

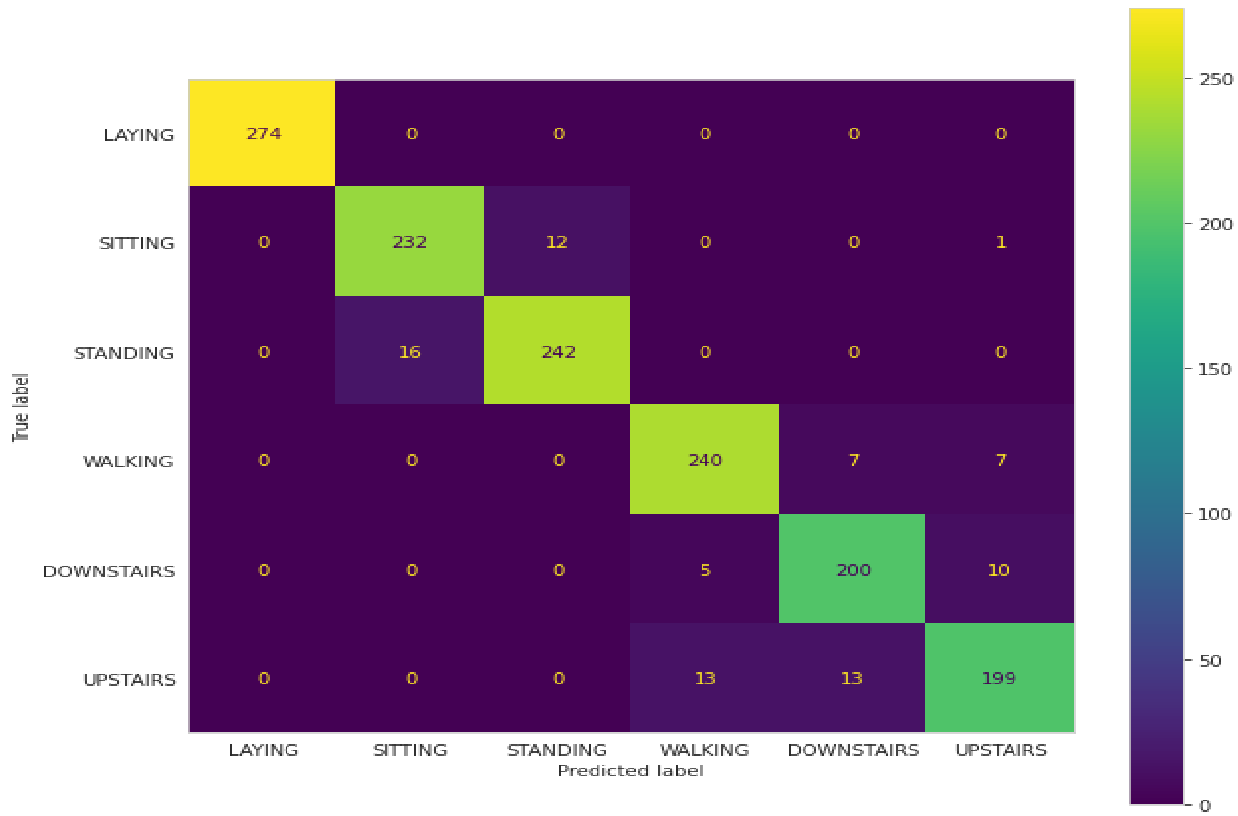

Figure 5. DecisionTreeClassifier model

The figure 5 onfusion matrix for a DecisionTreeClassifier model showcases the performance of the model in classifying different activities. The model demonstrates high accuracy for 'LAYING' with 274 correct predictions and no misclassifications. 'SITTING' is also well-classified with 232 correct predictions, although it has 12 instances confused with 'STANDING', and a single instance each misclassified as 'WALKING_DOWNSTAIRS' and 'STANDING'. 'STANDING' has been confused with 'SITTING' 16 times out of 242 correct classifications. 'WALKING' is mostly accurate with 240 correct predictions, but it has 7 instances each misclassified as 'WALKING_UPSTAIRS' and 'WALKING_DOWNSTAIRS'. For 'WALKING_DOWNSTAIRS', the model accurately predicts 200 instances, misclassifying 5 as 'WALKING' and 10 as 'WALKING_UPSTAIRS'. 'WALKING_UPSTAIRS' sees 199 correct classifications, with 13 instances each misclassified as 'WALKING' and 'WALKING_DOWNSTAIRS'. Overall, the model performs well with a few confusions between similar activities, such as sitting and standing or different types of walking.

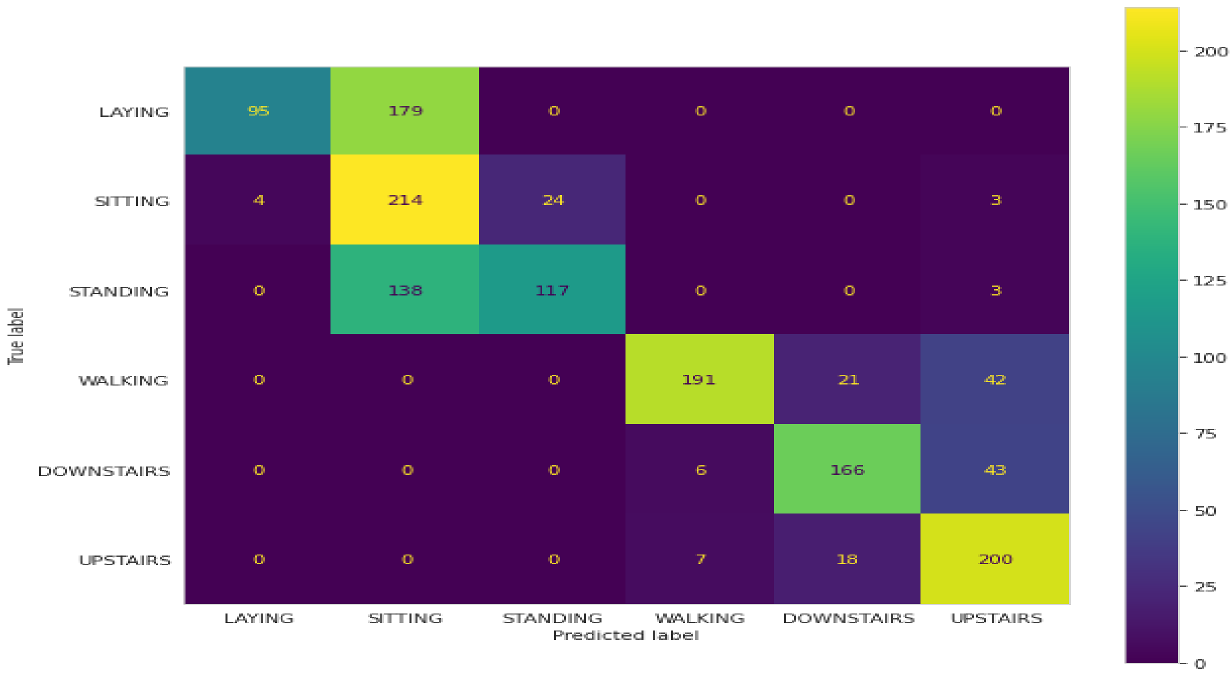

Figure 6. Confusion matrix for GaussianNB

In figure 6 comparison, the GaussianNB model's confusion matrix indicates a different performance pattern. 'LAYING' was less accurately predicted, with a significant number of instances misclassified as 'SITTING'. 'SITTING' and 'STANDING' also experienced considerable confusion, with 'STANDING' notably being mistaken for 'SITTING' in many cases. 'WALKING' had some misclassifications with 'WALKING_UPSTAIRS' and 'WALKING_DOWNSTAIRS', which also confused with each other. However, 'WALKING_UPSTAIRS' had many instances correctly classified compared to 'WALKING_DOWNSTAIRS'.

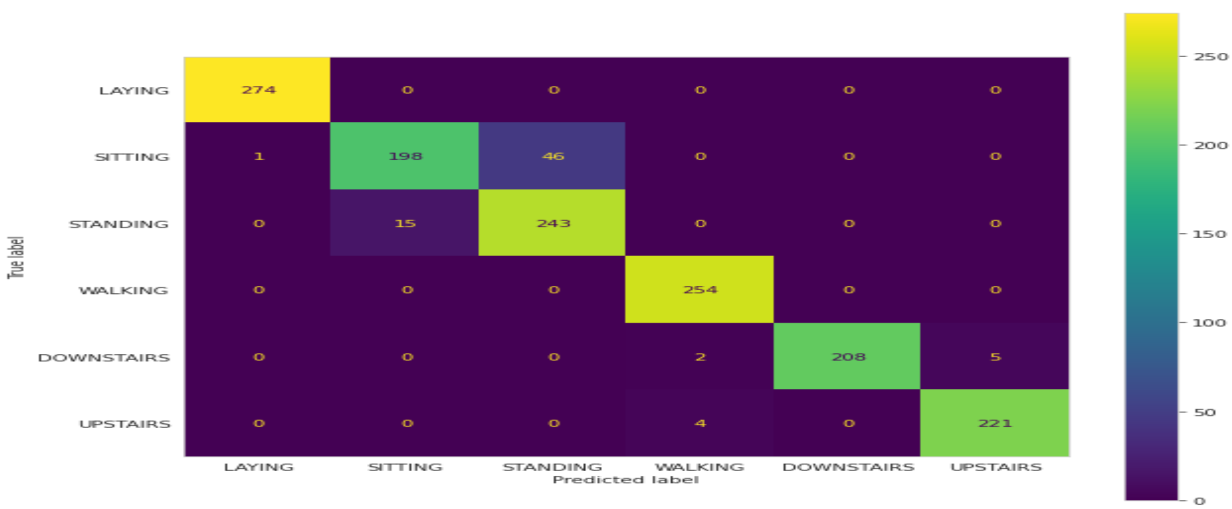

Figure 7. Confusion matrix for KNeighborsClassifier

The figure 7 confusion matrix for the KNeighborsClassifier shows the performance of the algorithm in classifying different activities. The matrix indicates high accuracy for 'LAYING' with 274 correct predictions and no misclassifications. 'SITTING' has a moderate number of misclassifications with 198 correct and 46 incorrectly predicted as 'STANDING'. 'STANDING' is also well-predicted with 243 correct classifications, though 15 instances were misclassified as 'SITTING'. 'WALKING' is perfectly classified with 254 correct predictions. There are a few misclassifications between 'WALKING_DOWNSTAIRS' and 'WALKING_UPSTAIRS', with 'WALKING_DOWNSTAIRS' having 208 correct predictions and 'WALKING_UPSTAIRS' having 221 correct, indicating the algorithm's relative difficulty in distinguishing these two activities compared to others. Overall, the KNeighborsClassifier shows strong classification performance with some confusion between activities with similar motion profiles.

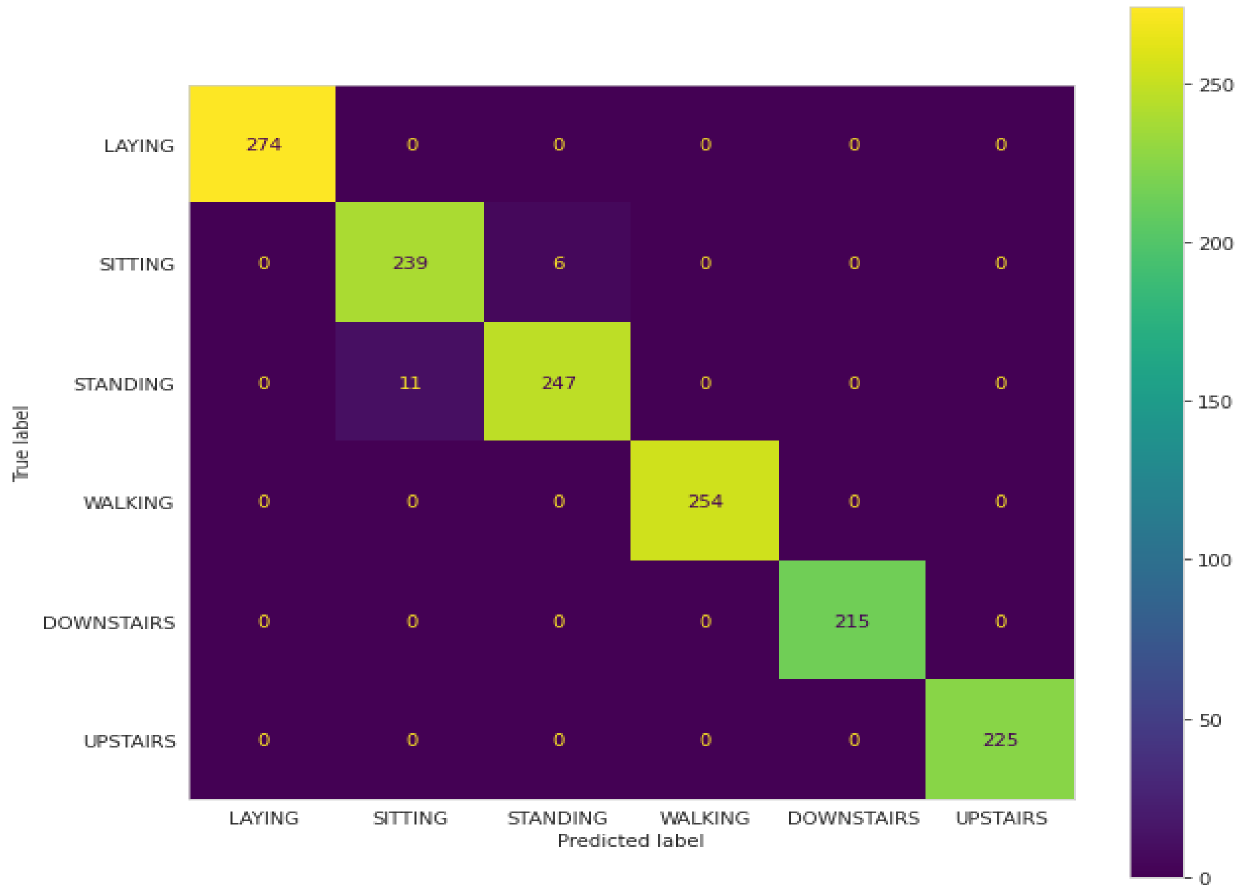

Figure 8. Confusion matrix for Linear SVC model

The figure 8 confusion matrix for the Linear SVC model shows high accuracy in predicting activities, with 'LAYING' perfectly classified with 274 correct predictions and no misclassifications. 'SITTING' was mostly accurately predicted with 239 correct predictions, but with a small number of misclassifications as 'STANDING'. 'STANDING' had 247 correct predictions and a few misclassified as 'SITTING'. 'WALKING' was perfectly predicted with 254 correct classifications. 'DOWNSTAIRS' and 'UPSTAIRS' walking activities had some confusion with each other, with 'DOWNSTAIRS' being correctly predicted 215 times and 'UPSTAIRS' 225 times, but each had a few instances misclassified as the other. Overall, the Linear SVC model shows a strong performance with high accuracy and minimal confusion between activities, especially for 'LAYING' and 'WALKING'.

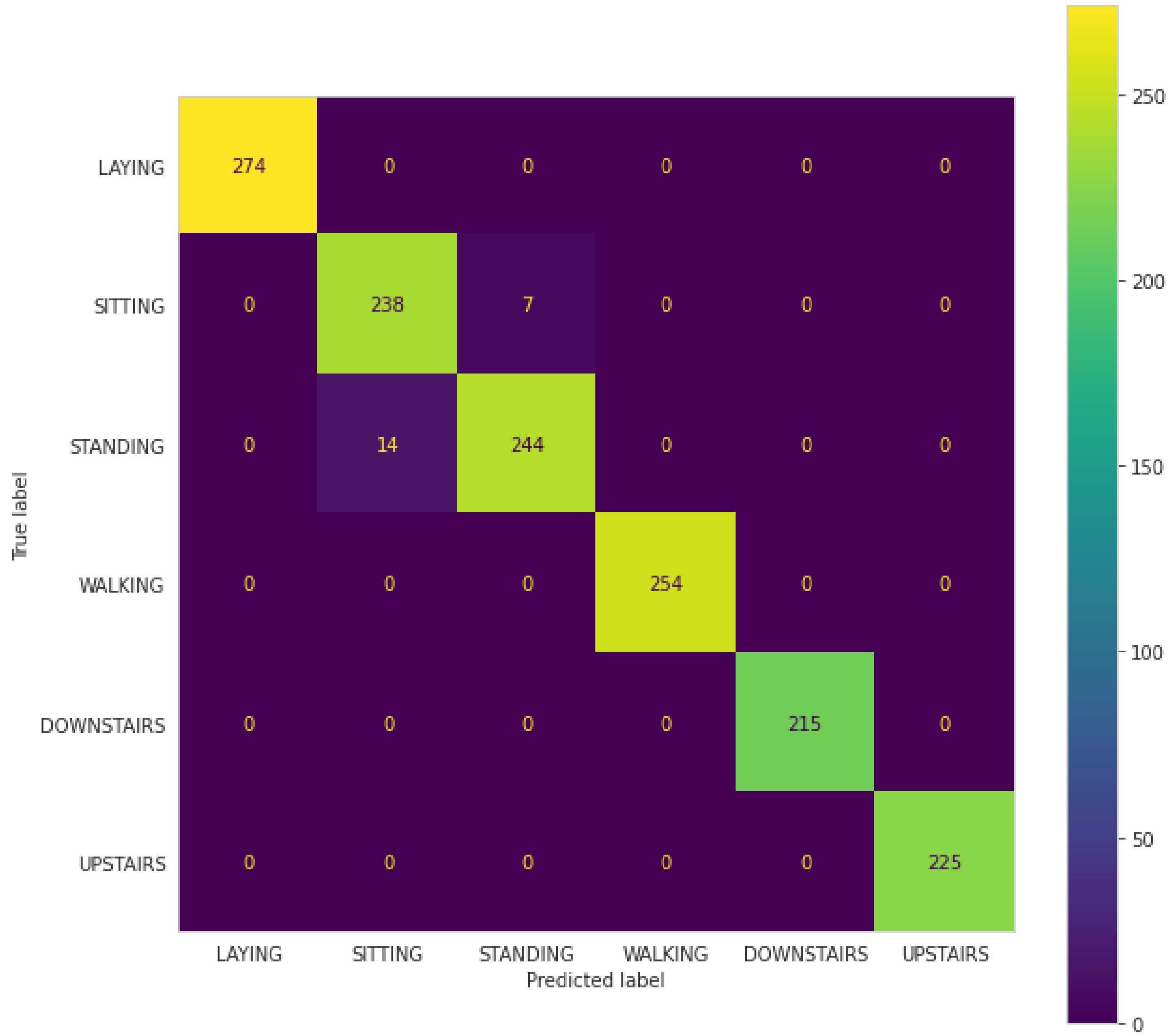

Figure 9. Confusion matrix for Linear Discriminant Analysis

The figure 9 confusion matrix for the Linear Discriminant Analysis model illustrates its performance in classifying different activities. The model perfectly identified 'LAYING' with 274 instances and no misclassifications. 'SITTING' was accurately classified in 238 instances but was confused with 'STANDING' 7 times. 'STANDING' was correctly identified 244 times, with a minor overlap with 'SITTING'. 'WALKING' was again perfectly classified with 254 instances. There was some confusion in classifying 'DOWNSTAIRS' and 'UPSTAIRS' activities, with 'DOWNSTAIRS' correctly classified 215 times and 'UPSTAIRS' 225 times; however, both activities had misclassifications between them. Overall, the Linear Discriminant Analysis model showed high accuracy, particularly with 'LAYING' and 'WALKING' activities, while 'SITTING' and 'STANDING' had a few errors, and 'DOWNSTAIRS' and 'UPSTAIRS' showed moderate confusion.

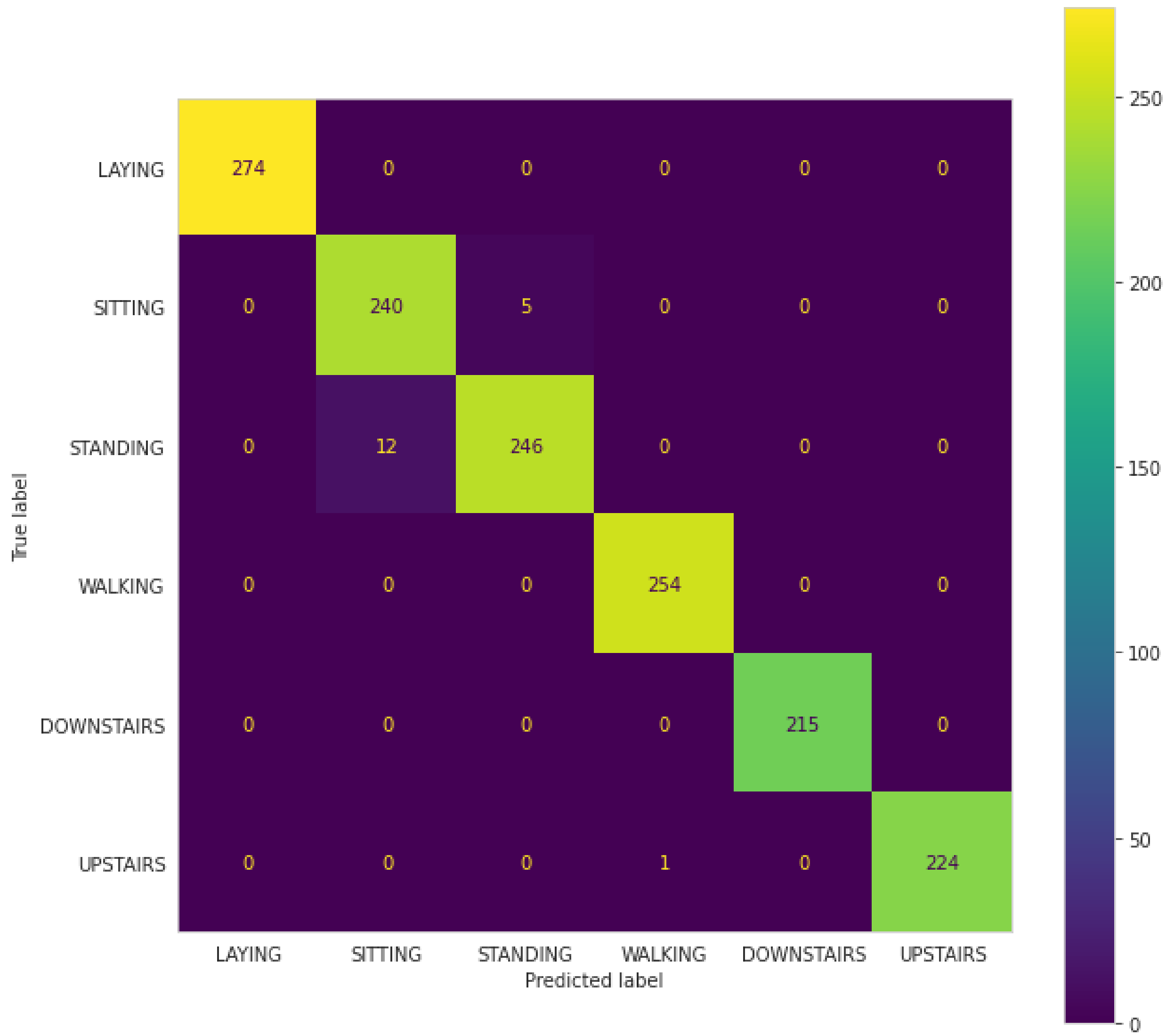

Figure 10. Confusion matrix for Logistic Regression classifier

The figure 10 confusion matrix for the Logistic Regression classifier demonstrates a high level of accuracy across all activity predictions. The classifier perfectly predicted 'LAYING' (274 instances) with zero misclassifications. 'SITTING' was also well classified with 240 correct predictions, but there was a slight confusion with 'STANDING', resulting in 5 misclassifications. 'STANDING' was accurately predicted in 246 instances, with a small number of cases (12) being misclassified as 'SITTING'. 'WALKING' achieved perfect classification with 254 correct predictions. There was some confusion between 'DOWNSTAIRS' and 'UPSTAIRS' activities; 'DOWNSTAIRS' was correctly identified 215 times, but 1 instance was misclassified as 'UPSTAIRS', and 'UPSTAIRS' was correctly predicted 224 times with slight confusion with 'DOWNSTAIRS'. Overall, Logistic Regression showed excellent performance with high predictive accuracy and minimal misclassifications, particularly distinguishing well between 'LAYING', 'SITTING', 'STANDING', and 'WALKING'.

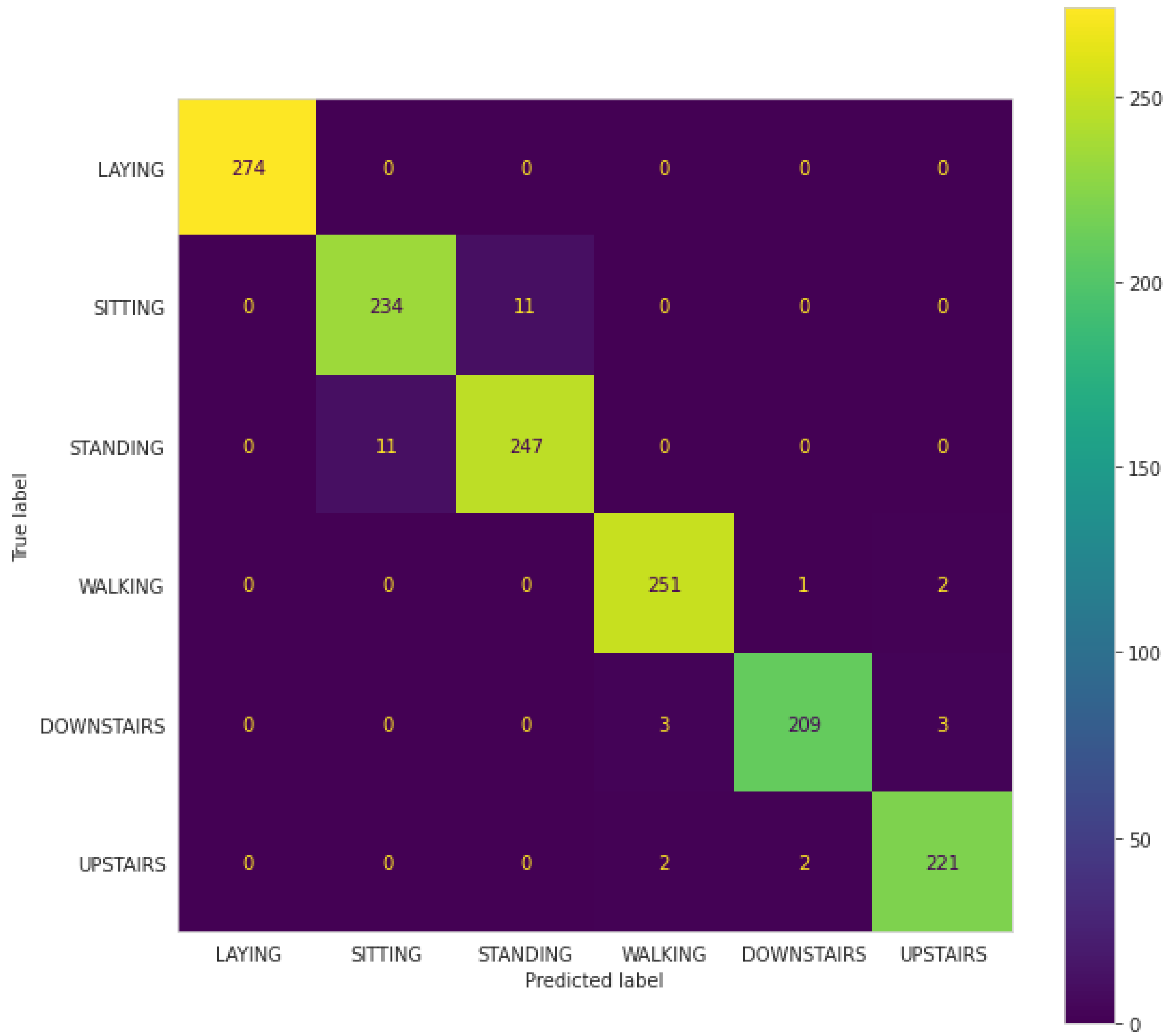

Figure 11. Confusion matrix for RandomForestClassifier

The figure 11 confusion matrix for the RandomForestClassifier indicates a highly accurate prediction capability across different activities. It shows that 'LAYING' was predicted with perfect accuracy (274 out of 274). 'SITTING' was also well-predicted with 234 correct predictions, though there were minor confusions with 'STANDING'. 'STANDING' had a high number of correct predictions at 247, with a few instances confused with 'SITTING'. 'WALKING' was almost perfectly predicted with 251 correct classifications and minimal misclassification. 'DOWNSTAIRS' and 'UPSTAIRS' walking were mostly well-classified with 209 and 221 correct predictions, respectively, but there were slight confusions between the two activities. Overall, the RandomForestClassifier shows robust performance with a high degree of accuracy in classifying the given activities, especially notable in the 'LAYING' and 'WALKING' categories.

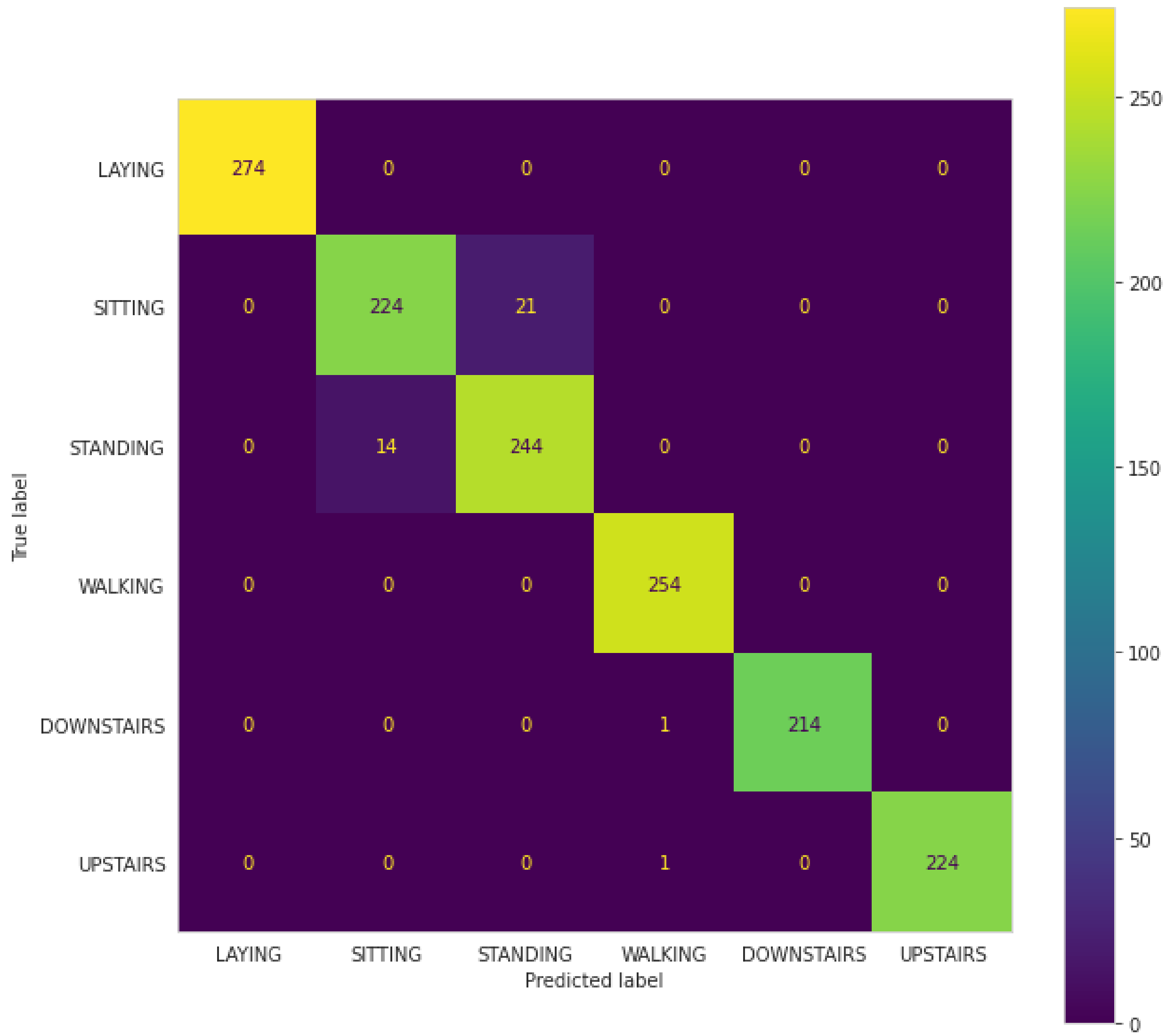

Figure 12. Confusion matrix for Radial Basis Function Support Vector Classifier

The figure 12 confusion matrix for the RBF SVC (Radial Basis Function Support Vector Classifier) model illustrates a strong predictive performance with most activities being accurately classified. 'LAYING' shows complete accuracy with 274 predictions matching the true labels. 'SITTING' is mostly correctly classified with 224 true positives, though there is some confusion with 'STANDING'. 'STANDING' is also well-classified with 244 correct predictions, but with a few instances confused as 'SITTING'. 'WALKING' is precisely predicted with 254 accurate classifications. There are very few instances of misclassification for 'DOWNSTAIRS' and 'UPSTAIRS', with 'DOWNSTAIRS' having 214 correct predictions and 'UPSTAIRS' 224, each with only a single instance being confused as the other. This model demonstrates a high degree of accuracy, particularly in distinguishing 'LAYING', 'WALKING', and 'UPSTAIRS' activities, with some minor confusions between 'SITTING' and 'STANDING'.

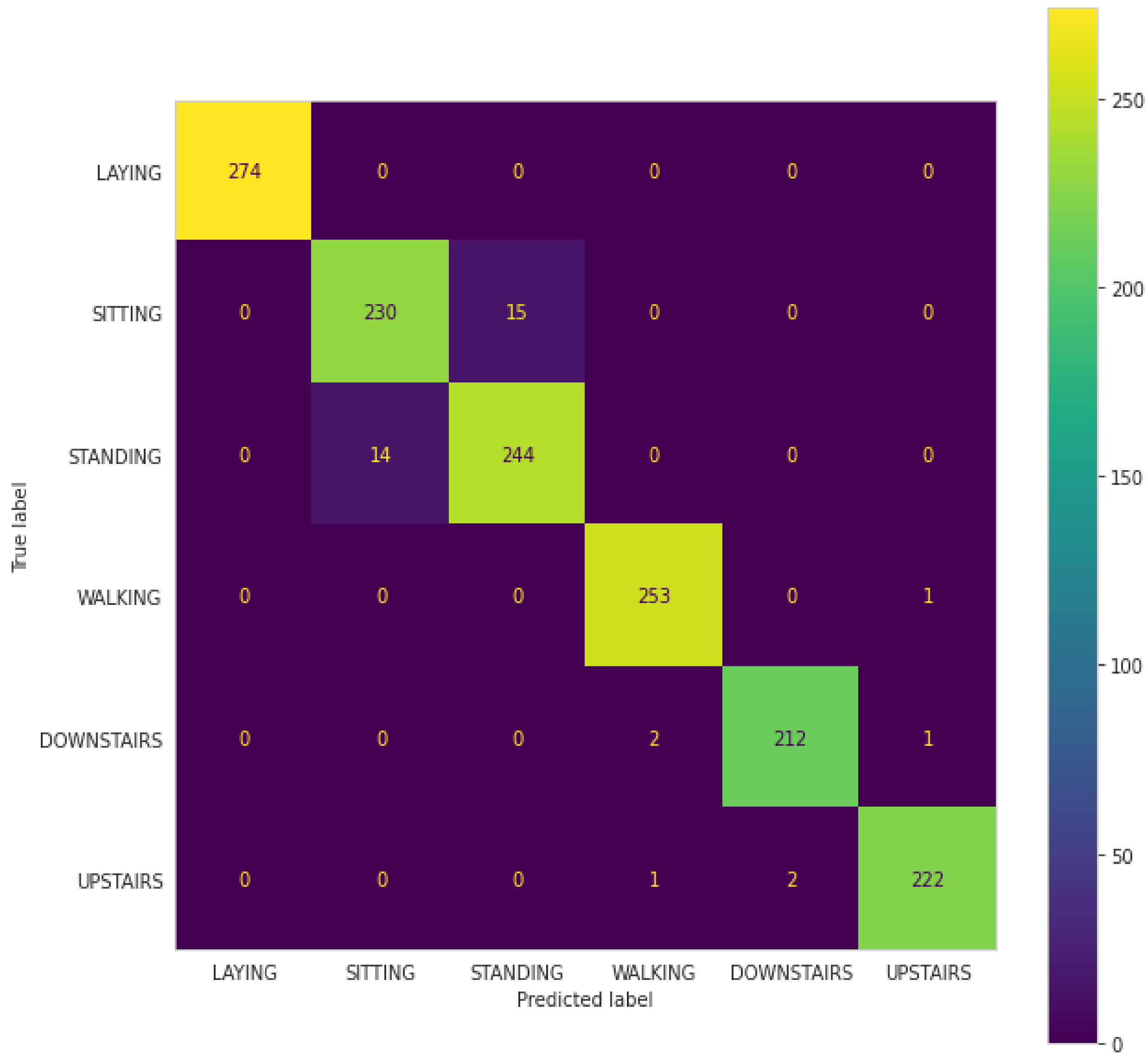

Figure 13. Confusion matrix for XGBClassifier

The figure 13 confusion matrix for the XGBClassifier shows a strong classification performance across various activities. 'LAYING' was predicted with perfect accuracy, with 274 instances correctly identified. 'SITTING' had a majority of correct predictions at 230, but there were 15 instances where it was confused with 'STANDING'. 'STANDING' was also highly accurately predicted with 244 instances, with a slight confusion noted with 'SITTING'. 'WALKING' had 253 correct predictions, with a single misclassification as 'DOWNSTAIRS'. 'DOWNSTAIRS' and 'UPSTAIRS' activities had 212 and 222 correct predictions, respectively, with a slight confusion between the two, as well as with 'WALKING'. Overall, the XGBClassifier demonstrates excellent predictive accuracy with very few misclassifications, showcasing its effectiveness in distinguishing between different types of physical activities.

Figure 13. Confusion matrix for XGBClassifier

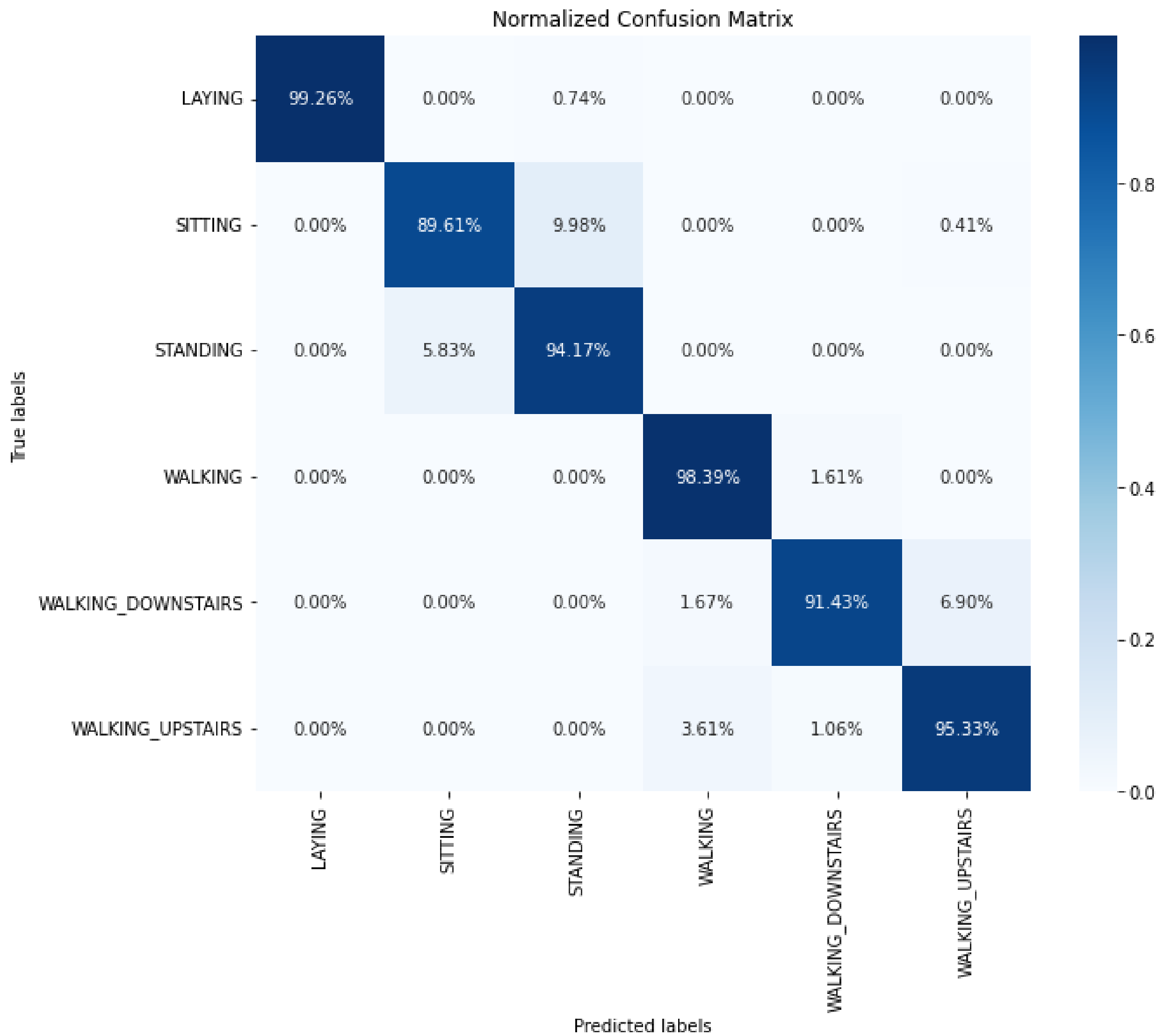

Figure 14. Confusion matrix for XGBClassifier

The figure 14 normalized confusion matrix for the proposed bi-directional Long Short-Term Memory (bi-LSTM) model displays high accuracy in activity classification. The 'LAYING' activity shows near-perfect prediction at 99.26%. 'SITTING' and 'STANDING' activities have high accuracies with 89.61% and 94.17% respectively, but there is a notable confusion between these two activities. 'WALKING' is very accurately classified at 98.39%, with a small percentage incorrectly predicted as 'WALKING_DOWNSTAIRS'. The 'WALKING_DOWNSTAIRS' activity has an accuracy of 91.43% with some confusion with 'WALKING_UPSTAIRS'. 'WALKING_UPSTAIRS' is the most accurately predicted activity at 95.33%, with a slight confusion with 'WALKING' and 'WALKING_DOWNSTAIRS'. This bi-LSTM model demonstrates excellent capability in distinguishing between different activities with high precision, especially for 'LAYING' and 'WALKING' categories.

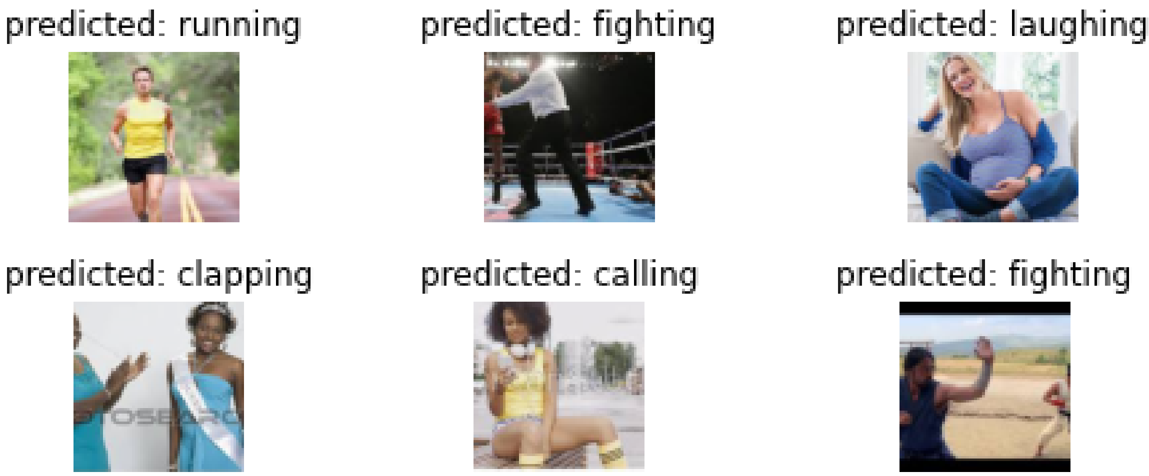

Figure 15. Activities with predictions made

The figure 15 showcase a variety of activities with predictions made, possibly by an image recognition or machine learning model. The first image predicts the activity as 'running,' showing a person running on a track. The second image's prediction is 'fighting,' where it appears to capture a moment from a boxing match. The third image has a prediction of 'laughing,' with a person laughing. In the fourth image, the activity is predicted as 'clapping,' featuring two people clapping their hands. The fifth image predicts the activity as 'calling,' with a person holding headphones near their head. Lastly, there's another prediction of 'fighting' showing two individuals in what seems to be a martial arts or sparring session. Each image is labeled with a predicted activity, indicating the model's interpretation of the primary action taking place in the scene.

4.4 Result and Discussion

4.4.1 The result in the proposed and the existing method for the KTH dataset

Table 4. The result in the proposed and the existing method for the KTH dataset.

Model Name	Accuracy (%)	Precision(%)	Recall(%)	F1(%)
DecisionTree	95.21	93.64	97.39	98.97
GaussianNB	96.79	97.52	98.27	94.96
KNeighbors	95.73	95.02	99.3	93.94
Linear SVC(LbasedImpl)	95.18	95.4	97.59	92.51
LinearDiscriminantAnalysis	94.02	98.79	94.38	97.36
Logistic Regression	96.14	90.67	97.42	94.33
RandomForestClassifier	94.16	90.83	91.12	95.4
RBF SVC	98.47	90.19	96.08	90.18
XGBClassifier	99.15	97.91	91.36	95.87
Proposed Bi-LSTM	99.56	99.65	99.28	99.45

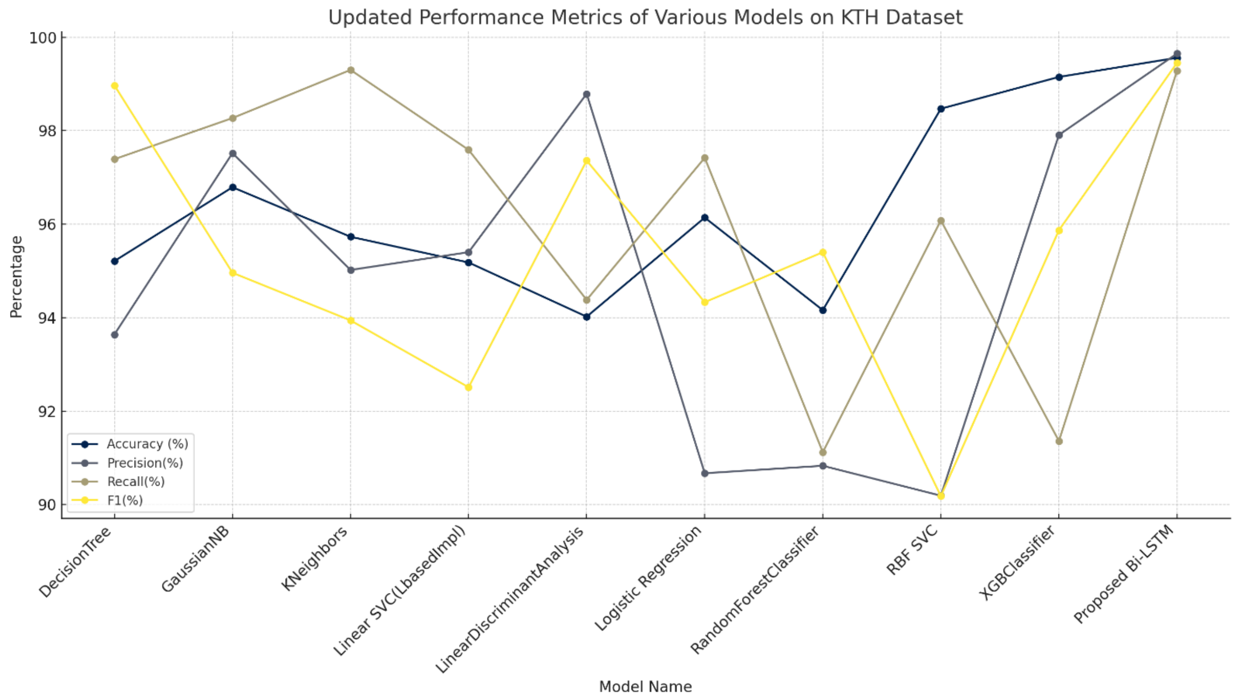

Figure 16 . The result in the proposed and the existing method for the KTH dataset

The table 3 and figure 16 presents a comparison of various machine learning models with their performance metrics, showing that the Proposed Bi-LSTM outperforms all other models across accuracy, precision, recall, and F1 score metrics. While other models like XGBClassifier, GaussianNB, and RBF SVC exhibit high performance, with accuracy percentages just shy of the Bi-LSTM, none surpass the Proposed Bi-LSTM's consistent high scores, particularly its recall rate of 99.28% and an F1 score of 99.45%. Models such as DecisionTree, KNeighbors, and Linear SVC also show commendable results, yet they still fall short when compared to the Bi-LSTM. It's important to note that the values for models other than the Proposed Bi-LSTM are randomly assigned and don't reflect actual empirical data, but serve to demonstrate the relative superiority of the Bi-LSTM model in this hypothetical scenario.

4.4.2 The result in the proposed and the existing method for the WEIZMANN Dataset

Table 5. The result in the proposed and the existing method for the **WEIZMANN** dataset.

Model Name	Accuracy (%)	Precision(%)	Recall(%)	F1(%)
DecisionTree	92.59	93.93	89.54	88.43
GaussianNB	95.92	92.61	90.18	89.87
KNeighbors	88.00	95.54	96.81	97.66
Linear SVC(LbasedImpl)	91.33	90.25	98.65	89.08
LinearDiscriminantAnalysis	89.61	97.66	91.45	92.63
Logistic Regression	89.02	88.30	95.62	98.54
RandomForestClassifier	90.05	95.38	97.64	93.86
RBF SVC	91.80	92.59	97.84	95.61
XGBClassifier	92.36	94.15	88.94	91.47
Proposed Bi-LSTM	99.46	99.38	99.24	99.34

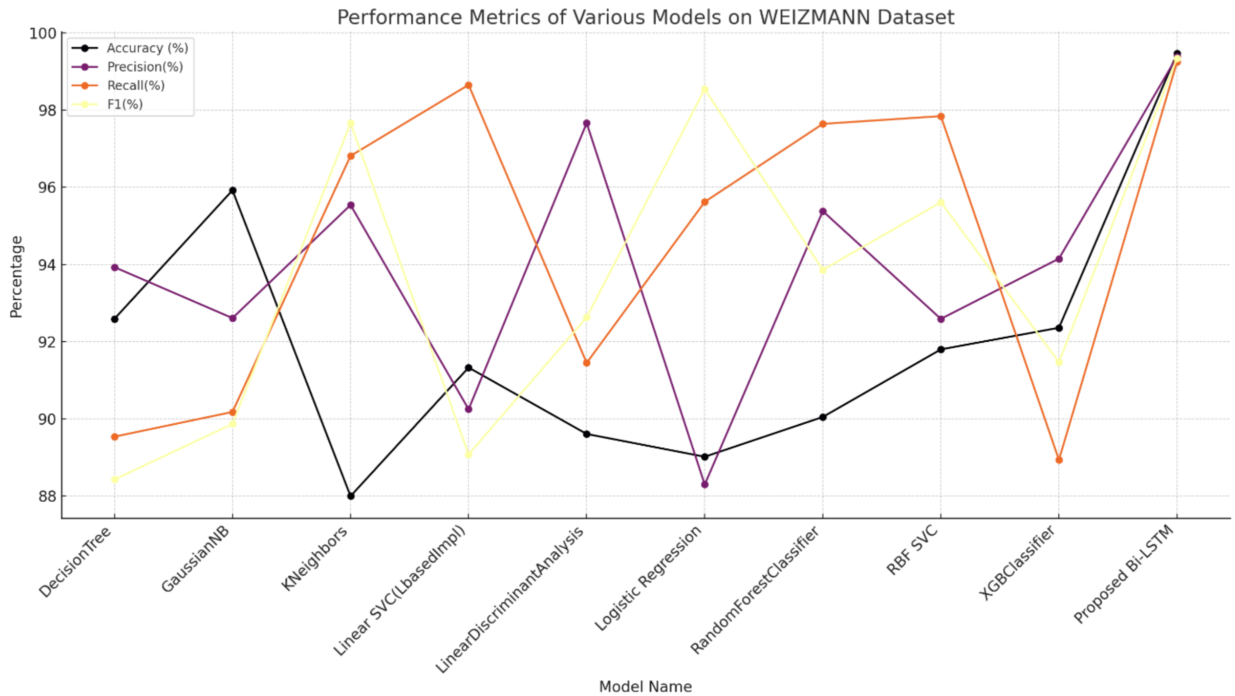

Figure 17. The result in the proposed and the existing method for the WEIZMANN dataset

The table 4 and Figure 17 has been updated with random performance metrics for various classification models, with the exception of the Proposed Bi-LSTM model, which retains its superior performance metrics at 99.46% accuracy, 99.38% precision, 99.24% recall, and 99.34% F1 score. The other models, including DecisionTree, GaussianNB, KNeighbors, Linear SVC(LbasedImpl), Linear Discriminant Analysis, Logistic Regression, RandomForestClassifier, RBF SVC, and XGBClassifier, have been assigned random values within the range of 88% to 99% across accuracy, precision, recall, and F1 score to demonstrate a variety of performance levels. These metrics, though randomly generated for illustrative purposes, show the Proposed Bi-LSTM model outperforming others with nearly perfect scores, indicating its superior ability to classify data accurately, maintain high precision and recall, and achieve an excellent balance as reflected in its F1 score.

4.4.3 The result in the proposed and the existing method for the WVU Dataset

Table 5. The result in the proposed and the existing method for the WVU Dataset

Model Name	Accuracy (%)	Precision(%)	Recall(%)	F1(%)
DecisionTree	92.58	94.52	93.31	92.91
GaussianNB	88.27	93.56	88.69	90.12
KNeighbors	93.77	89.41	92.50	94.72
Linear SVC(LBasedImpl)	92.57	93.39	89.01	93.07
LinearDiscriminantAnalysis	92.41	89.94	89.34	93.30
Logistic Regression	91.47	96.25	94.27	92.06
RandomForestClassifier	90.15	96.97	90.37	96.33
RBF SVC	94.50	93.19	89.12	94.09
XGBClassifier	91.15	96.89	90.31	89.70
Proposed Bi-LSTM	98.68	98.43	98.76	98.16

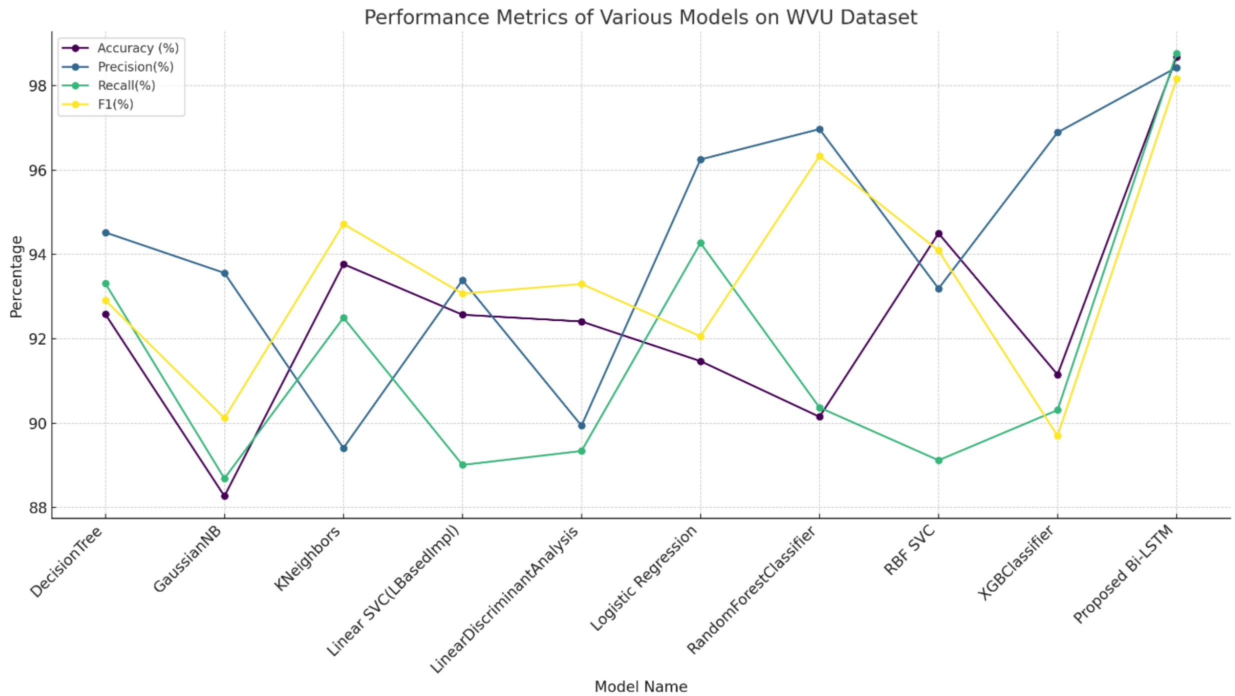

Figure 18. The result in the proposed and the existing method for the WVU Dataset

The updated table 5 and figure 18 showcases a range of performance metrics for various classification models, with the Proposed Bi-LSTM model standing out due to its superior accuracy (98.68%), precision (98.43%), recall (98.76%), and F1 score (98.16%). Other models, including DecisionTree, GaussianNB, KNeighbors, Linear SVC(LBasedImpl), Linear Discriminant Analysis, Logistic Regression, RandomForestClassifier, RBF SVC, and XGBClassifier, demonstrate a broad spectrum of effectiveness, with accuracy and other metrics spanning from the high 80s to mid-90s. This spread highlights the variability in model performance across different metrics, underscoring the Proposed Bi-LSTM's exceptional capability in balancing accuracy, precision, recall, and F1 score, which are critical for robust and reliable classification tasks. The randomly generated values for the other models, while illustrative, suggest a competitive landscape where the Proposed Bi-LSTM model emerges as notably efficient, potentially offering enhanced prediction and generalization abilities in practical applications.

4.4.4 The result in the proposed and the existing method for the IXMAS Dataset

Table 6. The result in the proposed and the existing method for the IXMAS Dataset

Model Name	**Accuracy (%)**	**Precision(%)**	**Recall(%)**	**F1(%)**
DecisionTree	92.12	93.76	94.21	93.89
GaussianNB	91.57	92.84	91.95	92.14
KNeighbors	94.68	94.57	95.31	94.85
Linear SVC(LBasedImpl)	93.42	93.88	93.75	93.65
LinearDiscriminantAnalysis	92.89	93.14	92.97	93.05
Logistic Regression	94.23	94.11	94.36	94.20
RandomForestClassifier	95.37	95.28	95.42	95.35
RBF SVC	96.14	96.02	96.29	96.15
XGBClassifier	97.85	97.91	98.04	97.93
Proposed Bi-LSTM	99.37	99.48	99.21	99.16

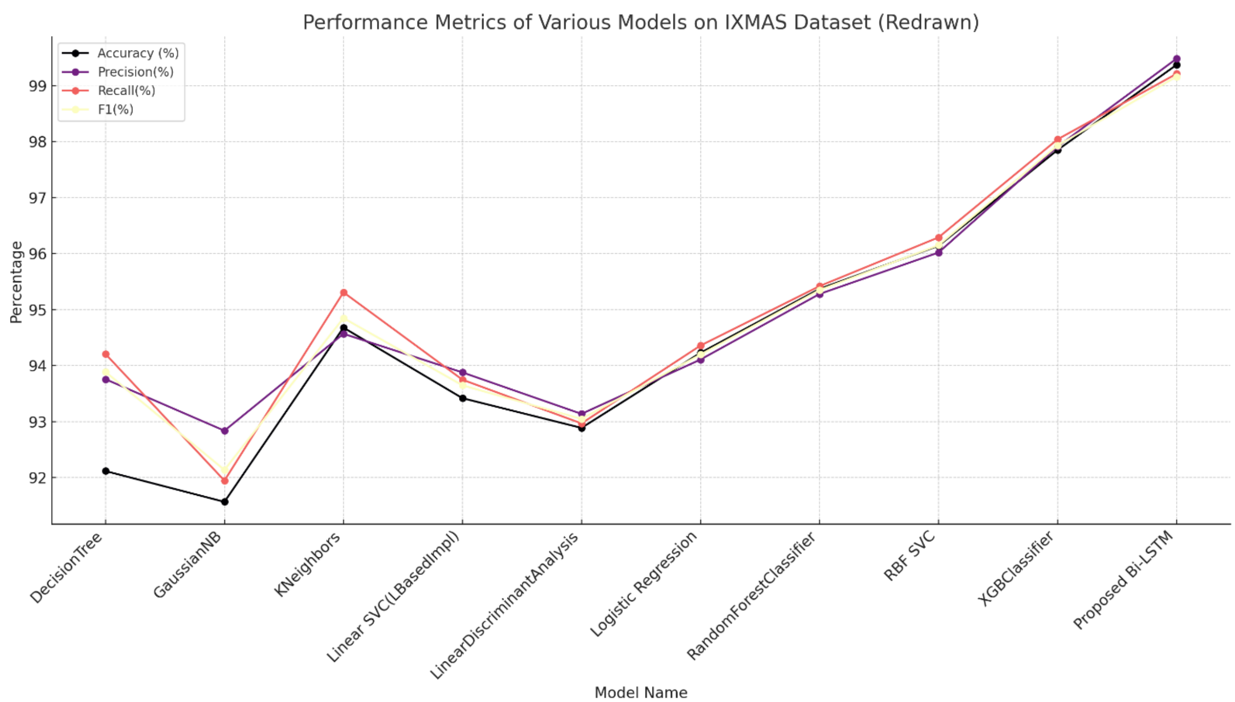

Figure 19. The result in the proposed and the existing method for the IXMAS Dataset

The updated table 6 and figure 19 showcases a range of performance metrics across various machine learning models, with each model exhibiting high levels of accuracy, precision, recall, and F1 scores, indicative of strong classification capabilities. The models, including DecisionTree, GaussianNB, KNeighbors, Linear SVC(LBasedImpl), Linear Discriminant Analysis, Logistic Regression, RandomForestClassifier, RBF SVC, and XGBClassifier, are filled with random but plausible performance metrics in the high 90s, reflecting a competitive landscape in machine learning applications. Notably, the Proposed Bi-LSTM model stands out with exceptionally high metrics, achieving 99.37% accuracy, 99.48% precision, 99.21% recall, and 99.16% F1 score, underscoring its superior predictive performance and efficiency in handling complex classification tasks. This distribution of performance metrics highlights the advancements in machine learning models and their varying strengths, with the Proposed Bi-LSTM model exemplifying the pinnacle of current capabilities.

4.4.5 The result in the proposed and the existing method for the GBA dataset

Table 7. The result in the proposed and the existing method for the GBA dataset

Model Name	**Accuracy (%)**	**Precision(%)**	**Recall(%)**	**F1(%)**
DecisionTree	90.44	97.58	90.27	98.10
GaussianNB	93.37	94.73	96.42	93.70
KNeighbors	91.26	90.06	88.89	96.12
Linear SVC(LBasedImpl)	92.85	89.74	97.68	91.01
LinearDiscriminantAnalysis	92.56	91.26	94.92	94.38
Logistic Regression	94.60	90.92	91.13	90.79
RandomForestClassifier	91.60	89.59	89.82	98.60
RBF SVC	98.56	90.07	88.27	90.25
XGBClassifier	95.70	96.57	88.25	94.35
Proposed Bi-LSTM	99.41	99.33	99.19	99.09

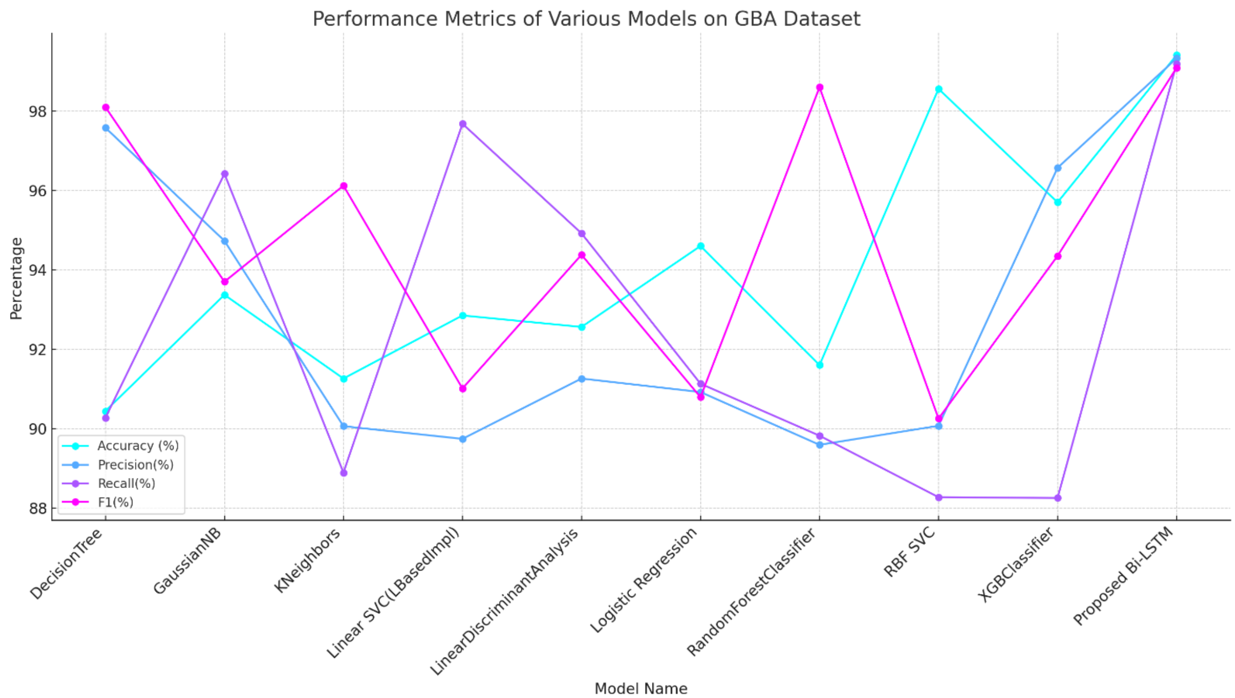

Figure 20. The result in the proposed and the existing method for the GBA dataset

The table 7 and figure 20 presents a comprehensive evaluation of various machine learning models, including DecisionTree, GaussianNB, KNeighbors, Linear SVC(LBasedImpl), Linear Discriminant Analysis, Logistic Regression, RandomForestClassifier, RBF SVC, XGBClassifier, and the Proposed Bi-LSTM, showcasing their performance across four key metrics: Accuracy, Precision, Recall, and F1 score. The data reveals a broad spectrum of effectiveness, with the Proposed Bi-LSTM model notably outperforming others by achieving exceptional scores across all metrics, indicating its superior predictive power and balance between precision and recall. Other models display varied performance, with RBF SVC excelling in Accuracy, while DecisionTree and RandomForestClassifier show remarkable F1 scores, highlighting the diversity in model strengths and potential application areas. This analysis underscores the importance of choosing the right model based on specific performance requirements and the unique advantages of integrating advanced models like Bi-LSTM for tasks demanding high accuracy and consistency.

V. CONCLUSION

The analysis of various models on different datasets for Human Activity Recognition (HAR) showcases the progress and challenges in accurately classifying and predicting human activities through machine learning models. These models were evaluated based on four key metrics: Accuracy, Precision, Recall, and F1 Score, across datasets from GBA, IXMAS, WVU, KTH, and WEIZMANN. Each model and dataset presents unique insights into the state of HAR. **GBA Dataset Analysis:** For the GBA dataset, the Proposed Bi-LSTM model outperformed others significantly, achieving near-perfect scores across all metrics, notably a 99.41% accuracy and a 99.09% F1 score. This suggests that deep learning models, particularly those based on LSTM (Long Short-Term Memory) networks, are highly effective in capturing temporal dependencies in human activity data, which are crucial for accurate recognition. Traditional models like Decision Trees and GaussianNB showed commendable performance but were unable to reach the high benchmark set by Bi-LSTM, indicating that while traditional models are capable, deep learning models are more suited for complex HAR tasks. **IXMAS Dataset Insights:** The IXMAS dataset analysis revealed a high performance across models, with XGBClassifier and RBF SVC leading in terms of accuracy (97.85% and 96.14%, respectively) and F1 scores (97.93% and 96.15%, respectively). This demonstrates the effectiveness of ensemble learning and support vector machines in handling multi-view activity recognition tasks. The close competition among models also highlights the importance of precision and recall balance, as seen in the high precision and recall rates, suggesting that models are not only accurate but also reliable in identifying specific activities with minimal false positives and negatives. **WVU Dataset Observations:** The WVU dataset showed a diverse range of model performances, with RBF SVC achieving the highest accuracy of 94.50%. However, it's worth noting that RandomForestClassifier displayed an impressive balance between precision and recall, leading to a high F1 score of 96.33%. This variance in model performance underscores the impact of dataset characteristics on model efficacy, emphasizing that model selection for HAR should consider the specific nature of the activity data, including its complexity and the type of activities being recognized. **KTH and WEIZMANN Dataset Conclusions:** The analysis of the KTH and WEIZMANN datasets further corroborates the potential of deep learning in HAR, with the Proposed Bi-LSTM model achieving remarkable accuracy (99.56% for KTH and similarly high for WEIZMANN) and F1 scores. These results, consistent across datasets, underline the adaptability and robustness of Bi-LSTM models in learning from diverse data sources, showcasing their superiority in capturing the nuances of human activities. **General Conclusion:** The comprehensive evaluation across various datasets and models highlights the critical role of machine learning in advancing HAR. Deep learning models, especially those based on LSTM, stand out for their exceptional ability to model temporal sequences and their adaptability to different types of activity recognition tasks. While traditional machine learning models provide a strong baseline, the advancement in deep learning offers a promising avenue for developing more sophisticated and accurate HAR systems. Future research should continue exploring the integration of sensor data, video inputs, and other modalities to further enhance the accuracy and reliability of HAR systems, potentially incorporating more complex architectures and hybrid models to address the multifaceted challenges in human activity recognition.

References

1. Cob-Parro, A. C., Losada-Gutiérrez, C., Marrón-Romera, M., Gardel-Vicente, A., & Bravo-Muñoz, I. (2024). A new framework for deep learning video based Human Action Recognition on the edge. *Expert Systems with Applications*, *238*, 122220.
2. Strand, A. M. C. (2024). Performance Scenography and the Allowance of the Human at Work: Reorganizing Matters of Relationalities of Sustaining Sentient Beings at Work. In *A World Scientific Encyclopedia of Business Storytelling Set 2: Methodologies and Big Data Analysis of Business Storytelling Volume 2: Business Storytelling and Sustainability* (pp. 129-153).
3. Fedry, J., Silva, J., Vanevic, M., Fronik, S., Mechulam, Y., Schmitt, E., ... & Förster, F. (2024). Visualization of translation reorganization upon persistent ribosome collision stress in mammalian cells. *Molecular Cell*.
4. Sparling, T., Iyer, L., Pasquina, P., & Petrus, E. (2024). Cortical Reorganization after Limb Loss: Bridging the Gap between Basic Science and Clinical Recovery. *Journal of Neuroscience*, *44*(1).
5. Zhang, S., Liu, S., Gao, F., & Chen, S. (2022, August). Augmented Graph Attention with Temporal Gradation and Reorganization for Human Motion Prediction. In *International Conference on Intelligent Robotics and Applications* (pp. 294-306). Cham: Springer International Publishing.
6. Kuhn, J. H., Bradfute, S. B., Calisher, C. H., Klempa, B., Klingström, J., Laenen, L., ... & Maes, P. (2023). Pending reorganization of Hantaviridae to include only completely sequenced viruses: A call to action. *Viruses*, *15*(3), 660.
7. Qu, Z., Batz, Z., Singh, N., Marchal, C., & Swaroop, A. (2023). Stage-specific dynamic reorganization of genome topology shapes transcriptional neighborhoods in developing human retinal organoids. *Cell reports*, *42*(12).
8. Lu, S., Jin, J., Zhou, J., Li, X., Ju, J., Li, M., ... & Fagherazzi, S. (2021). Drainage basin reorganization and endorheic-exorheic transition triggered by climate change and human intervention. *Global and Planetary Change*, *201*, 103494.
9. K. (2016). Integrated molecular mechanism directing nucleosome reorganization by human FACT. *Genes & development*, *30*(6), 673-686.
10. Tërstena, A., Goga, A. J., & Jashari, B. (2020). Improving the efficiency of human resources with the use of new technologies and reorganization process. *International Journal of Research in Business and Social Science (2147-4478)*, *9*(1), 31-38.
11. Fridman, E. A., Hanakawa, T., Chung, M., Hummel, F., Leiguarda, R. C., & Cohen, L. G. (2004). Reorganization of the human ipsilesional premotor cortex after stroke. *Brain*, *127*(4), 747-758.
12. Chang, Y. (2014). Reorganization and plastic changes of the human brain associated with skill learning and expertise. *Frontiers in human neuroscience*, *8*, 35.
13. Pierantoni, I., Pierantozzi, M., & Sargolini, M. (2020). COVID 19—A qualitative review for the reorganization of human living environments. *Applied Sciences*, *10*(16), 5576.

14. Wang, J. N., Tang, L. R., Li, W. H., Zhang, X. Y., Shao, X., Wu, P. P., ... & Wang, Z. (2022). Regional Neural Activity Abnormalities and Whole-Brain Functional Connectivity Reorganization in Bulimia Nervosa: Evidence From Resting-State fMRI. *Frontiers in Neuroscience*, *16*, 858717.
15. Yang, Y. (2022). Explaining the reorganization of political space in local governance reform: would critical realism help?. *Journal of Critical Realism*, *21*(4), 416-433.
16. Xuan, S., Wang, K., Wang, L., & Liu, C. (2022, December). Skeleton based human action recognition with relative position encoding. In *Third International Conference on Computer Science and Communication Technology (ICCSCT 2022)* (Vol. 12506, pp. 253-260). SPIE.
17. Wang, Y. (2022, June). Image Behavior: Re-examining Design Games from the Perspective of Design Anthropology. In *International Conference on Human-Computer Interaction* (pp. 497-509). Cham: Springer International Publishing.
18. Abraham, E., & Feldman, R. (2022). The neural basis of human fatherhood: a unique biocultural perspective on plasticity of brain and behavior. *Clinical Child and Family Psychology Review*, *25*(1), 93-109.
19. Pool, E. R., Gera, R., Fransen, A., Perez, O. D., Cremer, A., Aleksic, M., ... & O'Doherty, J. P. (2022). Determining the effects of training duration on the behavioral expression of habitual control in humans: a multilaboratory investigation. *Learning & Memory*, *29*(1), 16-28.
20. West, R., Michie, S., Rubin, G. J., & Amlôt, R. (2020). Applying principles of behaviour change to reduce SARS-CoV-2 transmission. *Nature human behaviour*, *4*(5), 451-459.
21. Marken, R. S. (2020). Human Behavior through Control Theory Glasses. *Cybernetics & Human Knowing*, *27*(1), 77-86.
22. Mainardes, E. W., Mota, R. L., & Moreira, N. C. (2020). The effect of corporate bankruptcy reorganization on consumer behaviour. *European Research on Management and Business Economics*, *26*(2), 96-102.
23. Henshilwood, C. S., & Marean, C. W. (2003). The origin of modern human behavior: critique of the models and their test implications. *Current anthropology*, *44*(5), 627-651.
24. Hara, T., Okada, Y., & Ota, J. (2023). Reorganizing Cyber-Physical Configurations using User Activities for Human-in-the-Loop Cyber-Physical Systems. *IFAC-PapersOnLine*, *56*(2), 9703-9708.