Efficient Human Activity Recognition through Multi-Sensor Data and Deep Learning Techniques

Rhenius A
CSE, KITS
Coimbatore, India.
rheniusa@karunya.edu.in

Anitha J
CSE, KITS
Coimbatore, India.
anitha_j@karunya.edu

Immanuel Alex Pandian S
ECE, KITS
Coimbatore, India.
immans@karunya.edu

Abstract* – Extensive research takes place in the field of Human Activity Recognition (HAR) systems with wearable computing that helps in healthcare applications, fitness tracking and smart environments. Sensor technology advances have propelled Human Activity Recognition systems to utilize innovative machine learning methods for accurate and robust activity classification. This work evaluates the effectiveness of four cutting-edge deep learning models—Temporal Convolutional Networks (TCN), Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN) and Gated Recurrent Units (GRU)—through their implementation on publicly accessible accelerometer and gyroscope HAR datasets. These classification models underwent testing through different activation functions namely ReLU, Leaky ReLU, and Swish to assess their practicality for classification operations. Standard evaluation metrics including accuracy, precision, recall, ROC-AUC and F1-score demonstrate the best features of each implemented model while providing descriptive visual representations. TCN delivered 98.59% accuracy and exhibited superior training speed efficiency thus establishing itself as the suitable AI algorithm for real-time implementation. This research offers an extensive breakdown for choosing optimal models in real-time HAR systems which supports advances in wearable computing and activity recognition technologies.

***Keywords* – Human Activity Recognition (HAR), Temporal Convolutional Networks (TCN), Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), Gated Recurrent Units (GRU), Multi-Sensor Data, Real-Time Applications.**

I. INTRODUCTION

The rapidly evolving Human Activity Recognition (HAR) discipline merges sensor technology with machine learning principles to detect human actions across healthcare applications and fitness monitoring alongside security and smart building domains [1]. HAR systems parse human movement patterns through sensor data analysis from accelerometers and gyroscope sensors to detect activities including walking, sitting, running, or lying down. Fast and precise activity identification capabilities create new opportunities for health monitoring systems, elderly care methods and automated intelligent systems.

The traditional methods for HAR required manual features to be extracted from sensor data together with classical machine learning (ML) models comprising Support Vector Machines (SVM) and K-Nearest Neighbors (KNN). These approaches led to general success rates yet remained unable to handle the complex mix of sensor data containing high dimensions and noise patterns. Traditional systems had limited ability to comprehend how sequential data develops chronologically even though this understanding is essential for detecting activities that show transitional characteristics or overlapping stages.

Deep learning (DL) transforms HAR systems by extracting features automatically and improving the representation of time sequences. Sequence data processing requires Long Short-Term Memory networks (LSTM) together with their variants Recurrent Neural Networks (RNN) and Gated Recurrent Units (GRU) for effective learning of temporal dependencies in real-world applications. The unique parallel processing abilities and stable gradient propagation mechanism of Temporal Convolutional Networks (TCN) continues to enhance HAR systems by solving computational efficiency and scalability problems.

Several challenges still prevent progress in human activity recognition tasks. Noise levels in sensor data along with activity differentiation and computational resource enhancement continue to present major obstacles for researchers. The application of attention mechanisms provides essential solutions to current challenges since they enable automatic coalescence of significant input data regions for better interpretations and performance improvements. The learning capacity and convergence speeds of models depend heavily on the selection of activation functions between ReLU, Swish and Leaky ReLU.

Human activity recognition activities are assessed by deep learning models including TCN, LSTM, RNN and GRU. The accelerometer and gyroscope data are available to researchers for both training and testing purposes. Multiple activation functions are utilized to evaluate models the assessment metrics include accuracy, recall, precision, F1-score and specificity.

The experimental results demonstrate that TCN works well for real-time applications along with achieving superior accuracy and testing performance. The combination of advanced activation functions alongside data fusion from multiple sensors generates a dependable solution for broad usage. The research develops dependable and efficient HAR systems that establish foundation for modern advancements in wearable computing and activity recognition technologies.

A. Literature Survey

Human Activity Recognition (HAR) has become a significant research area due to its applications in healthcare, smart environments, adaptive learning, and surveillance. HAR systems integrate ML and DL, and Internet of Things (IoT) technologies for enhancing activity recognition and addressing challenges such as data imbalance, computational constraints, and privacy concerns.

979-8-3315-0985-9/25/$31.00 ©2025 IEEE

Using smartphone sensors, HAR systems have demonstrated success in classifying physical activities and detecting falls. Accelerometers and gyroscopes have been instrumental in collecting high-resolution data, which is processed using ML models like KNN, SVM, and Random Forest (RF). SVM achieved 99.59% accuracy for walking activity classification, while KNN reached 85.71% accuracy in fall detection, emphasizing the reliability of ML for healthcare applications [1].

Research into hyperparameter tuning has further improved the performance of ML models. By applying grid search optimization to benchmark datasets like UCI, models such as SVM, RF, and Gradient Boosting achieved enhanced accuracy. Ensemble learning techniques has demonstrated robustness in handling diverse datasets, making them suitable for elder care and patient monitoring [2].

Real-world HAR systems have been advanced by leveraging data collected in natural environments. Studies using tree-based models, such as RF, showcased significant improvements in classification accuracy, increasing from 74.39% to 92.97%. This progress highlights the feasibility of deploying HAR systems in uncontrolled settings, bridging the gap between laboratory and real-world applications [3].

HAR benefits from ML applications demonstrated by the VanKasteren dataset to provide successful assistance for people with dementia-related cognitive impairments. The research utilized Locally Weighted Learning (LWL) and Multi-Layer Perceptron (MLP) algorithms to reach outstanding accuracy measurements of 98.81% and precision of 100% for assistive technology applications [4].

Imbalanced datasets, a common challenge in HAR, have been addressed through mixed-kernel-based weighted extreme learning machines (MK-WELM). Economical methods and kernel optimization improved the classification of minority-class activities, crucial for applications like anomaly detection and behavior monitoring. These methods enhance the reliability of HAR systems in diverse scenarios [5].

Emotion detection has also contributed to HAR advancements. By analyzing physiological signals such as EEG, ECG, and GSR, models like RF and KNN achieved remarkable accuracy. RF excelled in GSR and EEG recognition, while KNN performed well in ECG-based classification, demonstrating the potential to integrate behavioral and emotional contexts into HAR systems [6].

Hybrid deep learning models have become transformative in HAR research. Combining long short-term memory (LSTM) networks with one-dimensional convolutional neural networks (1D-CNN) achieved 98.42% accuracy on ultra-wideband (UWB) radar data. These models outperformed traditional methods and optimized energy efficiency, making them suitable for smart building systems and complex activity recognition [7].

Federated learning (FL) has addressed privacy and scalability concerns in HAR. FL distributes learning across edge devices, ensuring data privacy while maintaining accuracy. It has proven effective in healthcare and surveillance applications, where data sensitivity is a primary concern. This decentralized approach enables scalable and secure HAR solutions [8] [14].

Attention mechanisms have further refined HAR models by prioritizing relevant features and improving precision in complex tasks. These mechanisms are valuable for enhancing hybrid frameworks and multi-sensor systems, aligning with findings on the importance of feature extraction [9].

Edge computing has also played a significant role in HAR by reducing latency and energy consumption. Combined with FL strategies, edge computing ensures scalable and efficient HAR solutions, particularly in resource-constrained environments. This method is essential for real-world applications that demand swift and precise activity recognition [10].

Healthcare digital twins have been introduced as innovative frameworks in HAR. By creating digital replicas of wearable sensors, real-time data processing at edge devices reduces latency and improves prediction accuracy. This layered architecture enhances decision-making and classification performance, particularly for healthcare monitoring systems [11].

Transitional activity recognition has also gained attention. A deep learning framework merging 1D-Convolutional Neural Network (CNN) and LSTM achieved classification accuracies of 97.84% for transitions and 99.04% for dynamic activities. By focusing on postural transitions, this approach demonstrated the importance of incorporating temporal dynamics in HAR systems [12].

Multi-sensor fusion has advanced HAR through hybrid frameworks that combine 1D-CNN, bidirectional LSTM (BiLSTM), and attention mechanisms. These methods effectively extract spatial and temporal features, achieving competitive results across various sensor fusion patterns. This underscores the importance of integrating diverse data streams to improve recognition accuracy [13].

To tackle data heterogeneity, a personalized federated learning framework has been proposed. This approach calculates user similarity based on demographic attributes like age and gender, improving classification accuracy on datasets such as RealWorld and SisFall while preserving data privacy. Personalized models enhance adaptability in wearable sensor-based HAR systems [15].

In conclusion, the exhaustive literature in HAR addresses challenges like data imbalance, privacy concerns, and computational constraints through advancements in ML, DL, federated learning, and multi-sensor integration. Innovations such as hybrid frameworks, attention mechanisms, and edge computing enhance recognition accuracy and scalability, revolutionizing healthcare, smart environments, and adaptive learning systems.

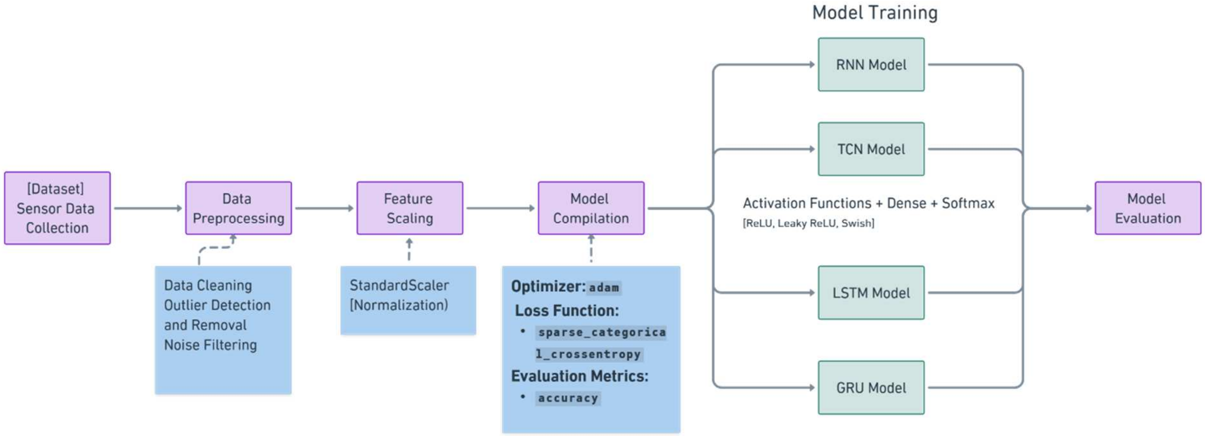

Fig.1 Proposed architecture for the HAR system

B. Our Contributions

Using multi-sensor datasets, this work evaluates four sophisticated deep learning models for Human Activity Recognition applications: Temporal Convolutional Networks (TCN), Long Short-Term Memory (LSTM) networks, Recurrent Neural Networks (RNN) and Gated Recurrent Units (GRU). This work's primary contributions are as follows:

1. Comprehensive Model Analysis: This research performs a detailed analysis of TCN, LSTM, RNN and GRU models performance by assessing training time and other performance metrics.
2. Effect of Activation Functions: Model performance is investigated based on different activation functions through analysis of ReLU and Leaky ReLU along with Swish.
3. Effective Testing and Training: The assessment prioritizes well-optimized training procedures together with optimal inference speeds to achieve real-time capability.

This research delivers critical knowledge required to develop fast HAR systems with high durability and accuracy.

II. MATERIALS AND METHODS

This work investigates TCN, LSTM, RNN and GRU models using HAR sensor data. Fig. 1 demonstrates the architecture of the proposed model.

A. Proposed Methodology

The proposed methodology for Human Activity Recognition (HAR) involves the following key steps:

1. **Data Acquisition:** A multi-sensor HAR dataset comprising accelerometer and gyroscope readings is utilized.
2. **Data Preprocessing:** Feature scaling, standardization, and train-test splitting are applied to ensure data consistency.
3. **Model Architectures:**
   - **TCN:** Focuses on hierarchical temporal feature extraction.
   - **RNN, LSTM, and GRU:** Capture sequential dependencies and address temporal complexities in data.
4. **Performance Evaluation:** The models' performance is assessed based on accuracy, precision, specificity, recall, F1-score, training time and inference time.

This approach ensures efficient temporal modeling, robust classification, and real-time applicability, addressing key HAR challenges such as noisy data and activity differentiation.

B. Pre-Processing Data

The preprocessing pipeline for the HAR dataset ensures clean, balanced, and normalized inputs for model training. The following steps are implemented:

1. **Data Cleaning**: It removes outliers using the Interquartile Range (IQR) method or address missing values.
2. **Feature Scaling**: StandardScaler is used to normalize features, but RobustScaler would handle outliers more effectively.
3. **Label Encoding**: LabelEncoder is correctly applied to prepare activity labels for machine learning algorithms.
4. **Train-Test Split**: The dataset is separated into training and testing sets using an 80-20 ratio to make sure an impartial evaluation.
5. **Data Reshaping**: Data reshaping using X[..., np.newaxis] ensures compatibility with deep learning models.

These steps ensure high-quality inputs for accurate HAR model training.

C. Activation mechanisms

Activation mechanisms are crucial in deep learning models as they introduce non-linearity by allowing the network to capture complex patterns.

1. **ReLU (Rectified Linear Unit):** Enhances sparsity and computational efficiency but may suffer from dying neurons.
2. **Leaky ReLU:** Addresses ReLU's limitations by allowing

small gradients for negative values, ensuring consistent gradient flow.
3. **Swish:** A smooth, non-linear activation that boosts convergence rates and performance in deep networks.

These activation functions are systematically evaluated in each model to identify the optimal configuration for HAR tasks.

D. Output Layer

The models' output layer comprises a fully connected dense layer followed by the Softmax activation function.

1. **Dense Layer:** Aggregates features extracted from previous layers and maps them to output classes.
2. **Softmax Activation:** Converts raw output logits into probability distributions, enabling multi-class classification.

This configuration ensures precise activity recognition, transforming complex sensor data into interpretable class labels.

E. Model Compilation and Training Configuration

The models are compiled with the following configurations:
- **Optimizer:** The Adam optimizer is chosen for its adaptability and efficient management of sparse gradients.
- **Loss Function:** Sparse categorical cross-entropy is used to measure difference between the predicted labels and actual labels.
- **Metrics:** Accuracy is chosen to evaluate classification performance.

Each model undergoes training for 20 epochs with a batch size of 32, promoting effective convergence and generalization.

F. Dataset

This study employs a publicly available HAR dataset, consisting of sensor data collected from accelerometers and gyroscopes. The dataset includes 10,299 samples with 561 features, representing six activity classes: walking (1722), standing (1777), sitting (1906), walking upstairs (1406), walking downstairs (1544) and lying down (1944). Data has been collected from 30 participants wearing smartphones, ensuring diversity in activity patterns. The dataset includes variations due to sensor placement and participant behavior, emphasizing the need for robust preprocessing and modeling techniques. This dataset serves as a valuable resource for training and assessing deep learning models in HAR tasks.

III. RESULTS AND DISCUSSION

This section evaluates TCN, LSTM, RNN, and GRU models for Human Activity Recognition, analyzing their performance and suitability for real-world applications. In this study, the Kaggle platform is utilized, and an NVIDIA Tesla P100 GPU is employed as the accelerator for model training.

The Fig. 2 shows the presence of outliers across various features, indicating inconsistency and potential noise and outliers of the dataset.

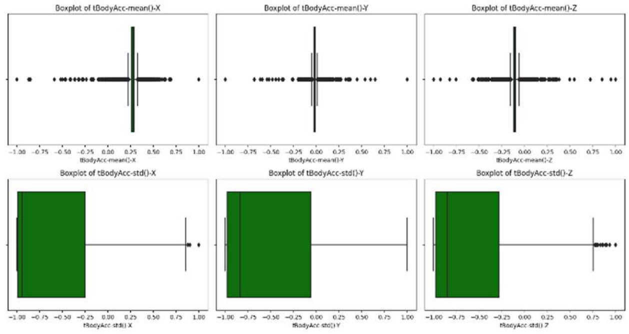

Fig. 2 The presence of outliers in the dataset

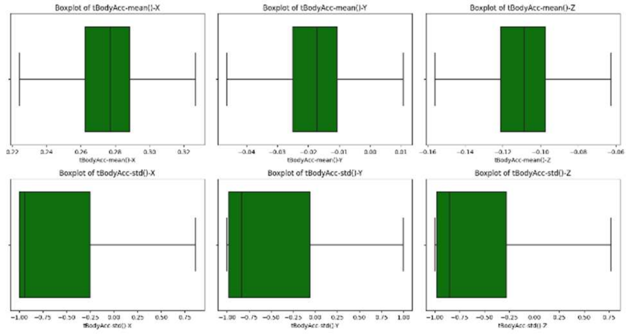

Fig. 3 Removal of outliers after the preprocessing

After preprocessing techniques like Z-score normalization and capping extreme values eliminated the outliers, improving data uniformity and ensuring cleaner inputs for model training as shown in the Fig. 3.

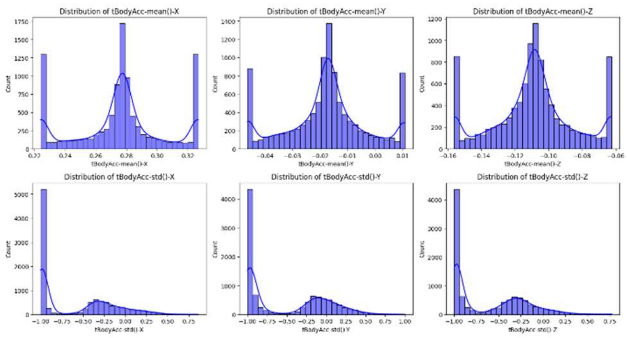

Fig. 4 Distribution of features in the dataset

The histograms in Fig. 4 illustrate the distribution of key features after post-preprocessing. Techniques like normalization and outlier handling are applied, resulting in a more uniform distribution. This ensures improved feature scaling and better alignment with the assumptions of the machine learning models.

The performance of Temporal Convolutional Networks (TCN), Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and Gated Recurrent Units (GRU) are evaluated on Human Activity Recognition (HAR) dataset using multiple activation functions: ReLU, Leaky ReLU, and Swish. The results highlight the unique strengths of each model, emphasizing their suitability for different applications.

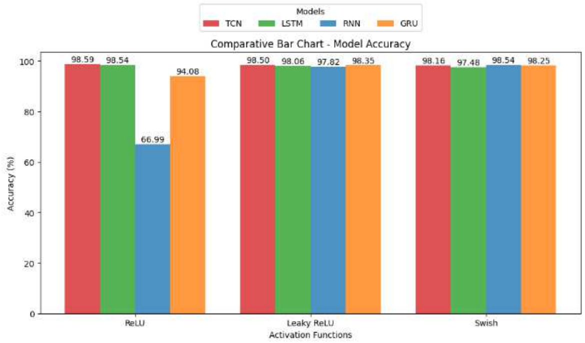

Fig. 5 Accuracy of the various models

This Fig.5 shows the accuracy of models, in which the TCN emerged as the most accurate model, achieving 98.59% with ReLU activation, followed closely by LSTM at 98.54% and GRU at 98.35%. TCN consistently outperformed other models across all metrics, demonstrating its ability to classify the activities effectively.

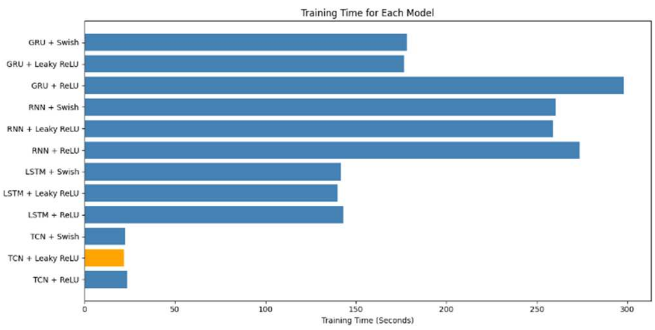

Fig. 6 Training time taken for each model

Fig.6 shows the training time efficiency where TCN exhibited the fastest training time, completing training in just 23.63 seconds with ReLU activation due to its parallel processing capability, whereas GRU and LSTM required significantly more time, averaging 195–298 seconds due to is sequential execution.

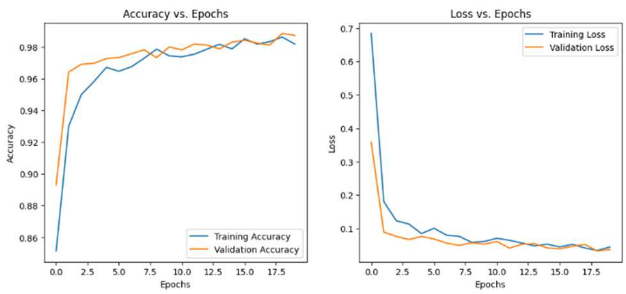

Fig. 7 Accuracy and loss epochs for the TCN model

The graph Fig.7 showcases the accuracy and loss trends for the TCN model across training epochs. The steady decline in loss and the corresponding rise in accuracy demonstrate the effective model convergence and robust learning for Human Activity Recognition tasks.

Key Observations:

1. **Model Performance:** TCN's ability to extract hierarchical temporal features through parallel processing, receptive field with dilated convolutions, backpropagation, not rely on recurrent connections, and GPU utilization made it particularly effective for HAR tasks.

2. **Activation Function Analysis:**

- **ReLU:** Showcased strong baseline performance across all models, offering simplicity and efficiency.
- **Leaky ReLU:** Improved gradient flow and reduced vanishing gradient issues, particularly for GRU models.
- **Swish:** Boosted accuracy for TCN and RNN, enhancing feature representation and convergence rates.

3. **Output Layers (Dense and Softmax):** The final dense layer mapped extracted features to activity classes, while the softmax activation ensured probabilistic outputs, enabling accurate and interpretable multi-class classification across all models.

The performance comparison in Table 1 highlights the effectiveness of TCN, LSTM, RNN, and GRU models with different activation functions (ReLU, Leaky ReLU, and Swish) for HAR. TCN consistently achieved high accuracy with minimal training time and inference time (the time required for the model to make predictions on new data), making it the most efficient. LSTM excelled in precision and recall, while RNN and GRU showed improved accuracy with Leaky ReLU and Swish. The results emphasize how the activation functions affect model performance and computational effectiveness.

Discussion:

The research demonstrates the selection of right combination of networks and activation approaches remains essential for HAR systems. TCN achieved both high accuracy rates and fast performance while LSTM and GRU showed strong competence in managing temporal relationships. TCN shows great potential for use in real-time HAR systems because it delivers fast performance while maintaining precise results. Future research will investigate how combining models and merging different data streams while utilizing edge computing will improve HAR system capabilities. Future work will focus on real-time metric coverage and real-world testing to validate the effectiveness of the models beyond controlled environments.

IV. CONCLUSION

The comprehensive study was conducted to identify the most effective deep learning systems for multi-sensor Human Activity Recognition by analyzing Temporal Convolutional Networks (TCN) alongside Long Short-Term Memory (LSTM) together with their variants Recurrent Neural Networks (RNN) and Gated Recurrent Units (GRU). The best performing deep learning model proven to be Temporal Convolutional Networks (TCN) which demonstrated 98.59% accuracy and 23.63 seconds training time using ReLU activation thus helping real-time applications.

The performance of LSTM and GRU models led to 98.54% accuracy for LSTM by maintaining excellent long-term dependence recognition capabilities although GRU models reached 98.35% accuracy advancing computational performance. Activation functions have proven crucial for model optimization which uses the activation functions such as ReLU, Leaky ReLU, and Swish to enhance the learning and performance of the deep learning models. This work demonstrates how optimized model designs and configurations help achieve HAR progress through wearable computing and activity recognition technology development.

Table.1 Model performance with various activation functions

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Time (Seconds)	Inference Time (Seconds)
TCN + ReLU	**98.59**	98.60	98.59	98.59	**23.63**	0.94
TCN + Leaky ReLU	98.50	98.51	98.50	98.50	**21.77**	0.70
TCN + Swish	98.16	98.19	98.16	98.15	**22.33**	0.75
LSTM + ReLU	98.54	98.60	98.54	98.54	142.88	0.83
LSTM + Leaky ReLU	98.06	98.16	98.06	98.06	139.81	0.81
LSTM + Swish	97.48	97.64	97.48	97.46	141.62	0.83
RNN + ReLU	66.99	69.49	66.99	65.37	273.47	0.94
RNN + Leaky ReLU	97.82	97.84	97.82	97.82	258.86	0.95
RNN + Swish	98.54	98.56	98.54	98.54	260.44	0.98
GRU + ReLU	94.08	94.07	94.08	94.07	298.14	1.68
GRU + Leaky ReLU	98.35	98.35	98.35	98.35	176.52	0.93
GRU + Swish	98.25	98.27	98.25	98.25	178.27	0.95

REFERENCES

[1] D. K. Agrawal, S. K. Udgata, and W. Usaha, "Leveraging Smartphone Sensor Data and Machine Learning Model for Human Activity Recognition and Fall Classification," *School of Telecommunication Engineering, Suranaree University of Technology, Nakhon Ratchasima, Thailand*, 2022.

[2] V. Ghate and S. Hemalatha, "A comprehensive comparison of machine learning approaches with hyper-parameter tuning for smartphone sensor-based human activity recognition," *School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India*, 2022.

[3] D. Garcia-Gonzalez, D. Rivero, E. Fernandez-Blanco, and M. R. Luaces, "New machine learning approaches for real-life human activity recognition using smartphone sensor-based data," *Department of Computer Science and Information Technologies, University of A Coruna, Spain*, 2023.

[4] "Machine Learning approach applied to Human Activity Recognition — An application to the VanKasteren dataset," *2023*.

[5] D. Wun, Z. Wang, Y. Chen, and H. Zhao, "Mixed-kernel-based weighted extreme learning machine for inertial sensor-based human activity recognition with imbalanced dataset," *Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, China*, 2022.

[6] A. Kumar and A. Kumar, "Human emotion recognition using Machine learning techniques based on physiological signals," *Department of Computer Science, Banaras Hindu University, India*, 2023.

[7] I. Khan, A. Guerrieri, G. Spezzano, and E. Serra, "A hybrid deep learning model for UWB radar-based human activity recognition," *University of Calabria, Italy*, *Boise State University, USA*, 2023.

[8] A. Trotta, F. Montori, L. Ciabattini, G. Billi, L. Bononi, and M. Di Felice, "Edge human activity recognition using federated learning on constrained devices," *University of Bologna, Italy*, 2023.

[9] P. P. Ariza-Colpas, R. P. Cuentas, and S. Butt-Aziz, "Semi-supervised ensemble learning for human activity recognition in CASAS Kyoto dataset," *2023*.

[10] J. Yang, Y. Xu, H. Cao, H. Zou, and L. Xie, "Deep learning and transfer learning for device-free human activity recognition: A survey," *School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore*, 2022.

[11] E. Bozkaya-Arasa, T. Onel, L. Eriskin, and M. Karatas, "Intelligent human activity recognition for healthcare digital twin," *National Defence University Turkish Naval Academy, Istanbul, Turkey*, 2024.<

[12] S. I. Khan, M. A. Khan, H. Dawood, G. F. Issa, and A. Hussain, "Transition-aware human activity recognition using an ensemble deep learning framework," *SZABIST University, Pakistan*, 2024.

[13] L. Zhang, J. Yu, Z. Gao, and Q. Ni, "A multi-channel hybrid deep learning framework for multi-sensor fusion enabled human activity recognition," *Donghua University, Shanghai, China*, 2024.

[14] D. Dayakaran and N. Kadiresan, "Federated Learning Framework for Human Activity Recognition Using Smartphones," *Amrita Vishwa Vidyapeetham, Coimbatore, India*, 2024.

[15] Y. Chai, H. Liu, H. Zhu, Y. Pan, and A. Zhou, "A profile similarity-based personalized federated learning method for wearable sensor-based human activity recognition," *Hefei University of Technology, China*, 2024.