science, engineering and health studies

Research Article

Science, Engineering and Health Studies
https://li01.tci-thaijo.org/index.php/sehs
ISSN (Online): 2630-0087

Enhancing human activity recognition with lightweight CNN models and integrated blocks

Teppakorn Sittiwanchai[1], Uttapon Khawnuan[2], and Nantakrit Yodpijit[2]*

[1] Center for Innovation in Human Factors Engineering and Ergonomics (CIHFE^2), Department of Production and Robotics Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand
[2] Center for Innovation in Human Factors Engineering and Ergonomics (CIHFE^2), Department of Industrial Engineering, Faculty of Engineering, King Mongkut's University of Technology North Bangkok, Bangkok 10800, Thailand

***Corresponding author**:*
Teppakorn Sittiwanchai
teppakorn.s@eng.kmutnb.ac.th

***Received**: 27 September 2023*
***Revised**: 1 February 2024*
***Accepted**: 2 February 2024*
***Published**: 2 September 2024*

***Citation**:*
Sittiwanchai, T., Khawnuan, U., and Yodpijit, N. (2024). Enhancing human activity recognition with lightweight CNN models and integrated blocks. Science, Engineering and Health Studies, 18, 24040001.

ABSTRACT

Human activity recognition (HAR) is crucial for health tracking, fitness monitoring, and fall detection systems. Recently, convolutional neural network (CNN) models have been proven to be highly effective for HAR tasks. This study aimed to enhance HAR performance by integrating specific architectural improvements, namely identity, convolutional, and bottleneck blocks, into lightweight CNN models. To evaluate the effectiveness of these enhancements, two data sets were utilized: HAR using smartphones data set version 1.0 (UCI-HAR) and wireless sensor data mining activity prediction data set version 1.1. The results indicated that the convolutional and identity block models outperformed the original lightweight CNN model on both data sets. The proposed models strike a balance between high performance and computational complexity, thereby making them suitable for real-world applications. The findings of this study contribute to the field of HAR and provide valuable insights for improving the recognition and classification of human activities.

Keywords: human activity recognition; convolutional neural networks; lightweight; identity block; convolutional block; bottleneck block

1. INTRODUCTION

Human activity recognition (HAR) plays a crucial role in health tracking, fitness monitoring, and fall detection systems (Gupta et al., 2022). It utilizes data such as acceleration, rotational speed, and location to accurately classify human actions (Wang et al., 2016), representing an important intersection between technology and human factors. Among the various machine-learning methods used in HAR, convolutional neural networks (CNNs) have demonstrated exceptional effectiveness (Phukan et al., 2022; Wang et al., 2019). Although recurrent neural networks (Murad and Pyun, 2017) and long short-term memory networks are effective in recognizing temporal patterns, they often fall short in HAR tasks owing to their extensive computational demands and less efficient spatial pattern learning than CNNs (Kashyap et al., 2022).

CNNs autonomously learn complex patterns from raw data and offer significant advantages for HAR (Straczkiewicz et al., 2021), eliminating the need for feature extraction (Souza et al., 2021), simplifying data preparation, and minimizing errors. In contrast to regular CNNs, lightweight CNNs are designed to balance model performance and computational efficiency (Zhou et al., 2020), making them suitable for resource-constrained applications, such as real-time or on-device implementations (Chen and Shen, 2017). MobileNet, a model often cited in HAR research for its lightweight design (Zhongkai et al., 2022), faces criticism for not achieving high performance, despite its suitability for high-end smartphones. Furthermore, onedimensional (1D) CNNs have demonstrated good performance in HAR tasks when customized properly (Kashyap et al., 2022; Xu et al., 2020).

The integration of block enhancements, such as identity, convolutional, and bottleneck blocks, can improve the CNN architecture (Barakbayeva and Demirci, 2023). The identity block enables the input from an early layer to bypass some layers and merge with the output of a later layer, mitigating gradient problems and facilitating the training of deeper networks (Negi et al., 2021). The convolutional block, comprising two convolutional layers with the same number of filters, enhances the model's capacity to discern complex data features, leading to improved activity differentiation (Agac and Incel, 2023). However, the bottleneck block alters the input dimensions before and after the convolution layers, aiding in computational demands and fostering the learning of complex data representations (Teng et al., 2021).

Although these integrated block enhancements have been extensively studied in resource-heavy CNNs, such as ResNet (Agac and Incel, 2023; Barakbayeva and Demirci, 2023; Negi et al., 2021; Ronald et al., 2021; Teng et al., 2021), limited research has been conducted on their effects on lightweight 1D CNNs. Therefore, this study aims to examine the effect of integrating these block enhancements into lightweight CNNs to enhance HAR performance tasks. By analyzing these effects, we aim to provide valuable insights that will advance the field of HAR. However, it is essential to enhance the recognition and classification of human activities, while also offering meaningful contributions to the broader field of HAR.

2. MATERIALS AND METHODS

2.1 Data sets

The models were trained and tested independently on two public HAR data sets: the activity recognition using smartphones data set version 1.0 (UCI-HAR) (Reyes-Ortiz et al., 2012) and the wireless sensor data mining (WISDM) activity prediction data set version 1.1 (Weiss, 2019). The UCI-HAR data set encompasses six activity classes (walking, going upstairs, going downstairs, sitting, standing, and lying down). The data set records the 3-axial linear acceleration and 3-axial angular velocity at a constant sampling rate of 50 Hz. Owing to its simplicity and thorough documentation, UCI-HAR has become a staple benchmark in HAR research (Ronald et al., 2021; Yin et al., 2022). In contrast, the WISDM data set exhibits different numbers of classes across its versions. This research used WISDM 1.1, which classified activities into six classes like those by UCI-HAR: walking, jogging, going upstairs, going downstairs, sitting, and standing). This dataset also captured the 3-axial linear acceleration and 3-axial angular velocity, albeit at a constant rate of 20 Hz. Training the models on these two data sets ensured a consistent evaluation of their performance across diverse data conditions.

2.2 Model architecture

The models were developed using Tensorflow with Keras on a computer equipped with an Intel® Core™ i7-6700 CPU running at 3.40 GHz, 16 GB RAM, and a GeForce GTX1060 GPU with 6 GB memory. The architectures of the proposed models are shown in Figure 1, with a detailed breakdown of the parameters listed in Table 1. The original lightweight CNN (2.2.1) (Ronald et al., 2021) utilized 1D convolutional layers and rectified linear unit (ReLU) activation functions, which served as a baseline for HAR tasks. Building upon this, we introduced three variants of lightweight CNNs, each integrating specialized blocks to enhance the performance. The first variant, a lightweight CNN with an identity block (2.2.2.1) (Negi et al., 2021) enhanced the original by enabling inputs from an early layer to bypass certain layers and merges with the output of a subsequent layer. The second variant, a lightweight CNN with a convolutional block (2.2.2.2) (Agac and Incel, 2023) incorporates a shortcut that includes a convolutional layer. Finally, the third variant, a lightweight CNN with a bottleneck block (2.2.2.3) (Teng et al., 2021) enhanced the original by introducing a multi-layered approach to feature extraction.

2.2.1 Original lightweight CNN

In this study, an original lightweight CNN model, specifically designed for HAR tasks (Ronald et al., 2021), was implemented. This model architecture comprised three convolutional sections, each featuring a 1D convolutional layer, batch normalization layer, and ReLU activation function. These sections were configured using 32, 64, and 32 filters respectively, each utilizing a kernel size of three. The 1D convolutional layer applied adjustable filters to process the input data and generate distinct feature maps. To prevent overfitting, a weight regularizer incorporating an L2 regularization factor of 0.0001 was applied. The inclusion of a batch normalization layer enhanced the model's performance, whereas the ReLU activation function introduced non-linearity into the system.

The processed output was then flattened into a 1D vector and relayed to a fully connected layer containing 100 neurons. This layer was responsible for classifying the extracted and down-sampled features. Another ReLU activation function processed the output before it progressed to the model's final stage. At this stage, the model featured a fully connected output layer equipped with a number of neurons equal to the output classes. It utilized a SoftMax activation function to produce a probability distribution across multiple classes, rendering it suitable for multiclass classification tasks.

The model used the Adam optimizer with a learning rate of 0.0005. It was constructed using categorical cross-entropy as the loss function and accuracy as the performance evaluation metric. The final output layer was designed for multiclass classification, rendering the model suitable for HAR tasks. The output ensured a probability distribution for each activity class, thereby delivering comprehensive and interpretable results.

2.2.2 Proposed lightweight CNNs with integrated blocks

The proposed lightweight CNN models with integrated blocks refine the architecture of the original lightweight CNN by introducing enhancements through distinct integrated blocks: the identity, the convolutional, and the bottleneck blocks. The development of these enhanced models stems from a series of trials and adjustments, drawing inspiration from more resource-intensive models, such as ResNet (Agac and Incel, 2023; Barakbayeva and Demirci, 2023; Negi et al., 2021; Ronald et al., 2021; Teng et al., 2021), which have previously demonstrated excellent performance in HAR tasks. To preserve the lightweight concept of the original model and facilitate comparisons, all the proposed models strategically integrated these blocks while keeping the number of layers as close as possible to the original architecture.

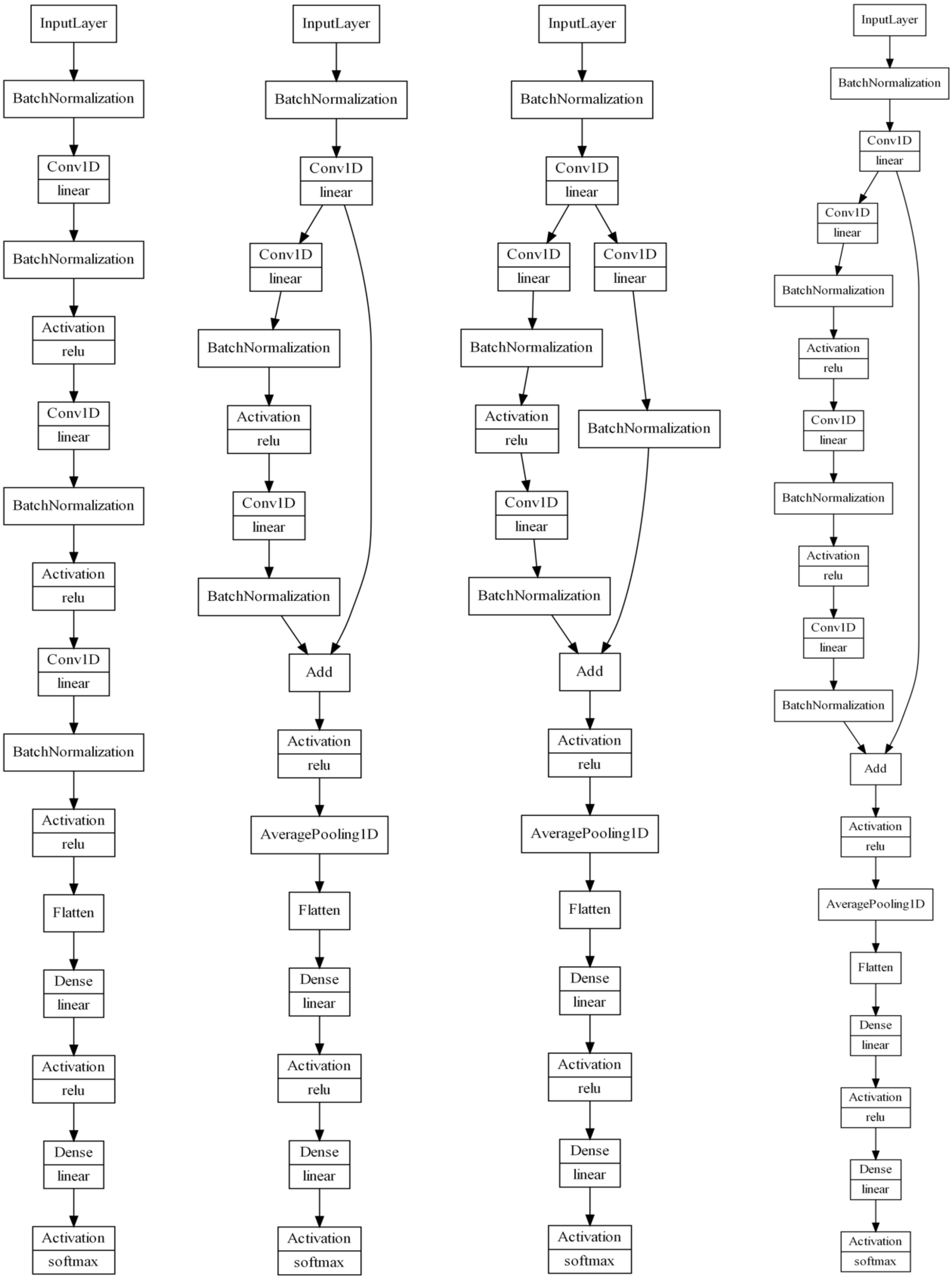

Figure 1. Model architecture: (a) original lightweight CNN (Ronald et al., 2021), (b) proposed lightweight CNN with an identity block, (c) proposed lightweight CNN with a convolutional block, and (d) proposed lightweight CNN with a bottleneck block

Table 1. Various parameters for models

Model	Total	Trainable	Non-trainable
(a) Original CNN (Ronald et al., 2021)	269,934	269,672	262
(b) Proposed lightweight CNN with identity block	90,702	90,568	134
(c) Proposed lightweight CNN with convolutional block	91,886	91,688	198
(d) Proposed lightweight CNN with bottleneck block	171,054	170,856	198

2.2.2.1 Proposed lightweight CNN with an identity block

The proposed lightweight CNN model, featuring an identity block, begins with a 1D convolutional layer with 32 filters and a kernel size of three, followed by an identity block. This block comprises two 1D convolutional layers, each equipped with 32 filters and a kernel size of three. A ReLU activation function and batch normalization were applied between the convolutional layers. This block includes a shortcut connection that directly adds the input to the output, thus mitigating the vanishing gradient problem that often arises in deeper networks. Diverging from the original lightweight CNN design, this model introduces an average pooling layer with a pool size of three precedes right before the flattening layer. This addition reduces the output size and summarizes the extracted features, whereas the remaining network architecture emulates the original lightweight CNN architecture.

2.2.2.2 Proposed lightweight CNN with a convolutional block

The proposed lightweight CNN with a convolutional block comprising a 1D convolutional layer utilizing 32 filters and a kernel size of three, followed by connection to a convolutional block. This block comprises two 1D convolutional layers, each with 32 filters and a kernel size of three. These layers were interspersed using a ReLU activation function and batch normalization. This model also included a shortcut path in the convolutional block. Here, the input is first processed through a 1D convolutional layer with 32 filters and a kernel size of one before it is added to the block output. An average pooling layer with a pool size of three precedes a flattening layer. The remaining network architecture closely mimicked the original lightweight CNN.

2.2.2.3 Proposed lightweight CNN with a bottleneck block

The proposed lightweight CNN with a bottleneck block comprises a 1D convolutional layer equipped with 64 filters and a kernel size of three, which then connects to a bottleneck block. This block consisted of three 1D convolutional layers. The first layer uses 16 filters (1/4 of 64) with a kernel size of one, the second layer uses 16 filters but expands the kernel size to three, and the third layer increases the number of filters to 64, maintaining a kernel size of one. Following each convolutional layer, a ReLU activation function was applied. This block features a shortcut connection that directly adds an input to the output. An average pooling layer with a pool size of three was initiated prior to the flattening layer. The remaining network architecture closely mimicked the original lightweight CNN.

2.3 Model evaluation

This study utilized a robust evaluation methodology for models, using a K-fold cross-validation approach (Ismail et al., 2023; Wang et al., 2023). This method involved dividing the data set into K subsets (k = 10 in this case), training the model on K-1 subsets, and evaluating it on the remaining subsets. This process was repeated K times, using a different subset as the test set.

Each model was compiled using the categorical cross-entropy loss function, which is a common choice for multiclass classification problems (Zhou et al., 2019). The Adam optimizer was selected for training the models because of its proven effectiveness and efficiency. The initial learning rate was set to 0.0005 (Ronald et al., 2021), and the models were trained for 1,000 epochs. The training process was performed using a batch size of 64 (Ronald et al., 2021).

To prevent overfitting and unnecessary training, the training process integrated an early stopping mechanism that halted training when no discernible improvement was observed in the model performance on the validation set for 100 consecutive epochs.

Moreover, a learning rate scheduler was employed to reduce the learning rate by a factor of 0.8 to a minimum threshold of 0.0001 (Ronald et al., 2021). This occurred when the model performance did not show significant improvement after 10 epochs, a condition determined by the patience parameter.

To capture and preserve the best model performance throughout the training process, the weights were saved at the end of each epoch. This method ensured that the optimal set of weights was retained, safeguarding against potential declines in the model performance in subsequent epochs. For data sets characterized by an imbalanced class distribution, class weights were calculated to adjust the loss function, signaling the model to account for the imbalance.

2.4 Performance metrics

Model effectiveness in HAR tasks was assessed using widely accepted performance metrics: accuracy, precision, recall, and F1-score. The accuracy of the model reflected the proportion of all correctly predicted observations. Precision measured the proportion of correctly predicted positive observations. Recall, also known as sensitivity, signified the ability of the model to correctly identify all actual positive cases. The F1-score combined precision and recall into a single measure by calculating their weighted average, offering a balanced perspective for both metrics (Wang et al., 2023).

These performance metrics were computed in two stages: initially, during each iteration of the K-fold cross-validation process, the metrics were calculated using the validation set of the current fold. This process provides a measure of the model's performance during the training phase. Following the completion of all the cross-validation folds, the metrics were assessed using an unseen test set. This step provided an unbiased evaluation of the performance of the fully trained model.

3. RESULTS

3.1 UCI-HAR

The analysis of the model performance on the UCI-HAR data set revealed insightful findings. Figures 2–9 show the changes in accuracy and loss across epochs for each of the four models, including the original lightweight CNN and three proposed lightweight CNNs with integrated blocks. Table 2 outlines the average performance metrics for each model on the UCI-HAR data set, sorted by the F1 score achieved during testing. The proposed lightweight CNN with a convolutional block yielded the highest F1 score of 0.9625 for testing and 0.9927 for validation, slightly outperforming the original lightweight CNN scores of 0.9623 (testing) and 0.9912 (validation). This pattern was consistent across other metrics, such as accuracy, recall, and precision, suggesting that the convolutional block model slightly improved the model's performance. Following closely was the proposed lightweight CNN with an identity block, which exhibited F1 scores of 0.9609 (testing) and 0.9927 (validation). The proposed lightweight CNN with a bottleneck block demonstrated the lowest performance, with F1 scores of 0.9557 (testing) and 0.9924 (validation). Despite being the lowest, these scores still indicate a relatively high and robust performance level. Figure 10 shows the best confusion matrix for fold 8 of the proposed lightweight CNN with a convolutional block on the UCI-HAR data set, providing insights into true positive, false positive, true negative, and false negative predictions. Additionally, Figures 11 and 12 depict the accuracy vs. epoch and loss vs. epoch plots, respectively, for the same fold and model on the UCI-HAR data set.

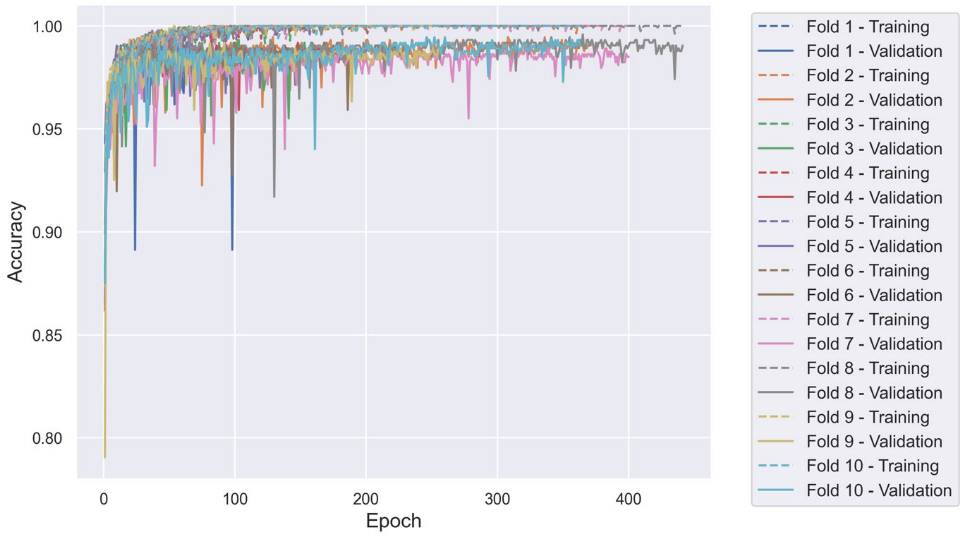

Figure 2. Accuracy vs. epoch plot of the original lightweight CNN on the UCI-HAR data set

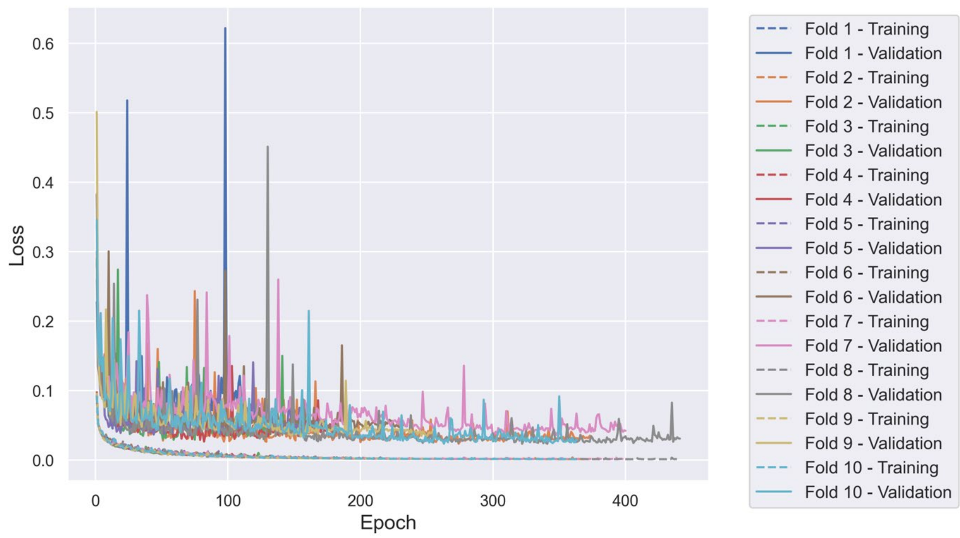

Figure 3. Loss vs. epoch plot of the original lightweight CNN on the UCI-HAR data set

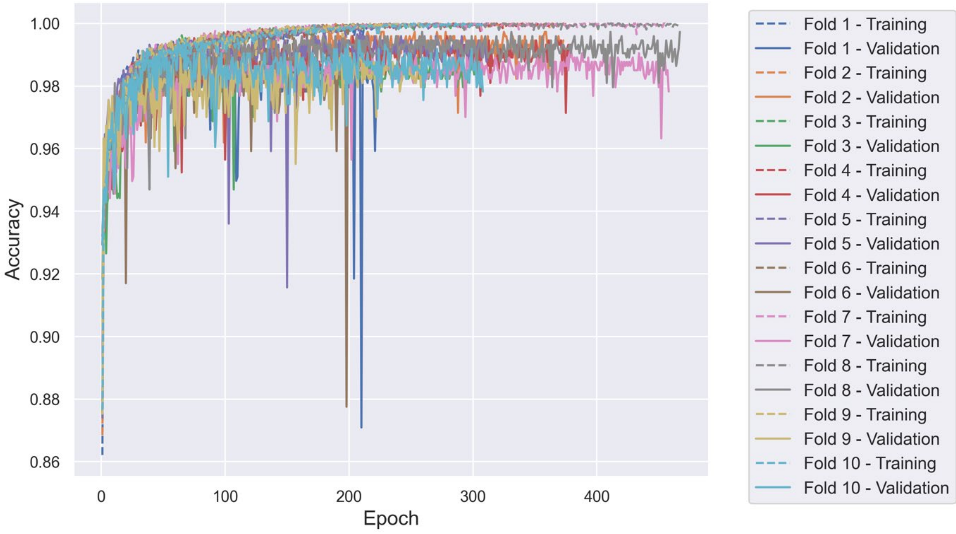

Figure 4. Accuracy vs. epoch plot of the proposed lightweight CNN with an identity block on the UCI-HAR data set

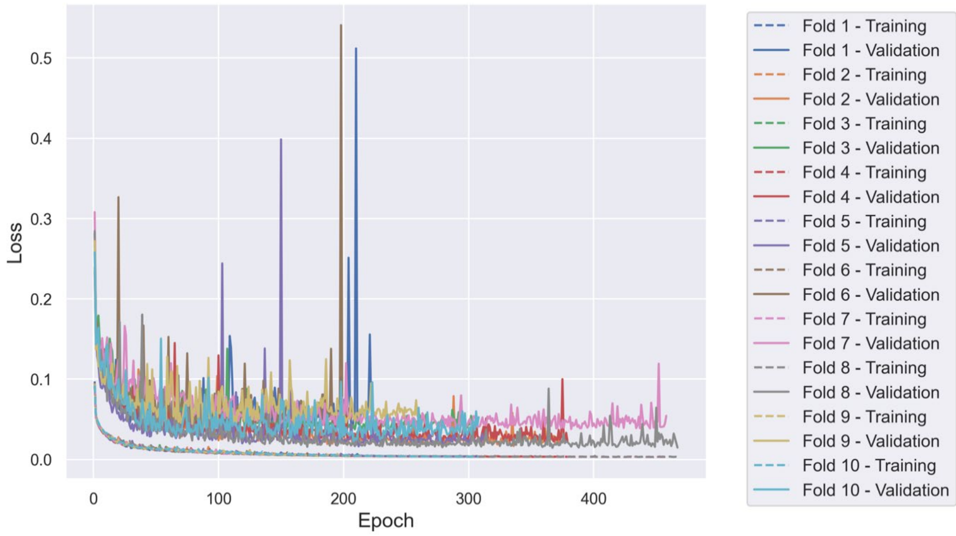

Figure 5. Loss vs. epoch plot of the proposed lightweight CNN with an identity block on the UCI-HAR data set

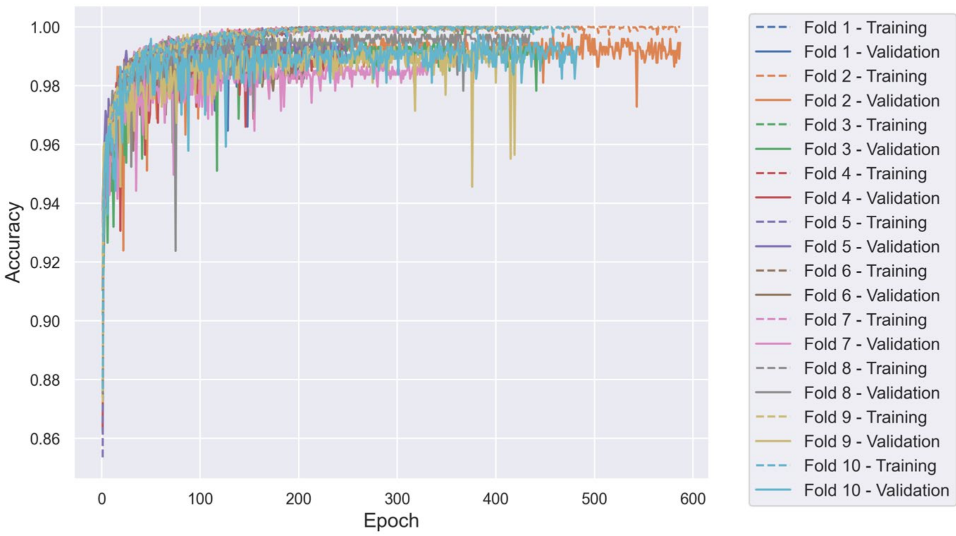

Figure 6. Accuracy vs. epoch plot of the proposed lightweight CNN with a convolutional block on the UCI-HAR data set

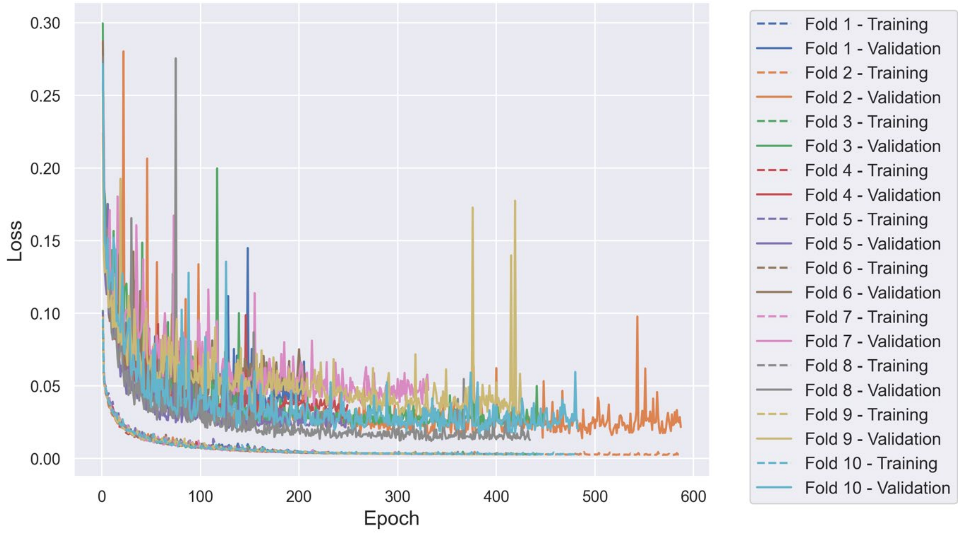

Figure 7. Loss vs. epoch plot of the proposed lightweight CNN with a convolutional block on the UCI-HAR data set

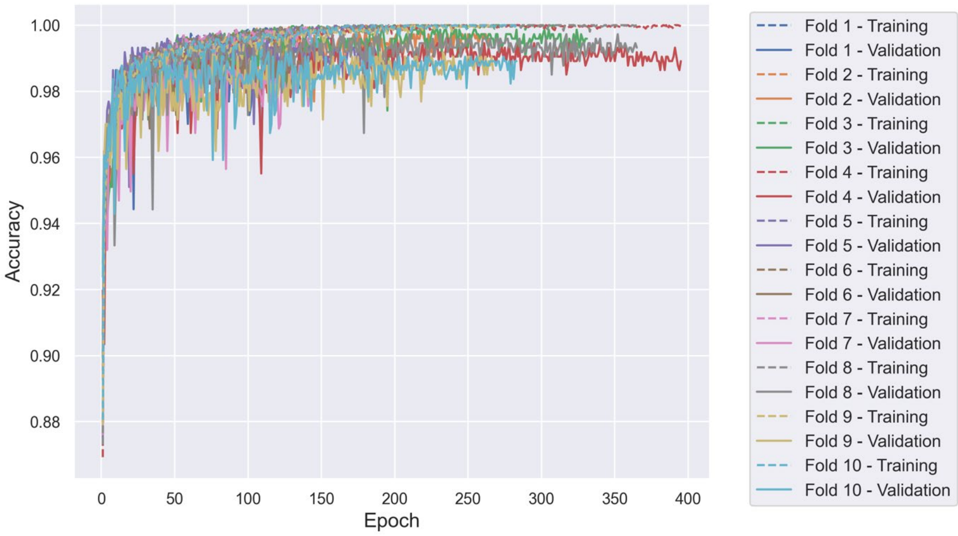

Figure 8. Accuracy vs. epoch plot of the proposed lightweight CNN with a bottleneck block on the UCI-HAR data set

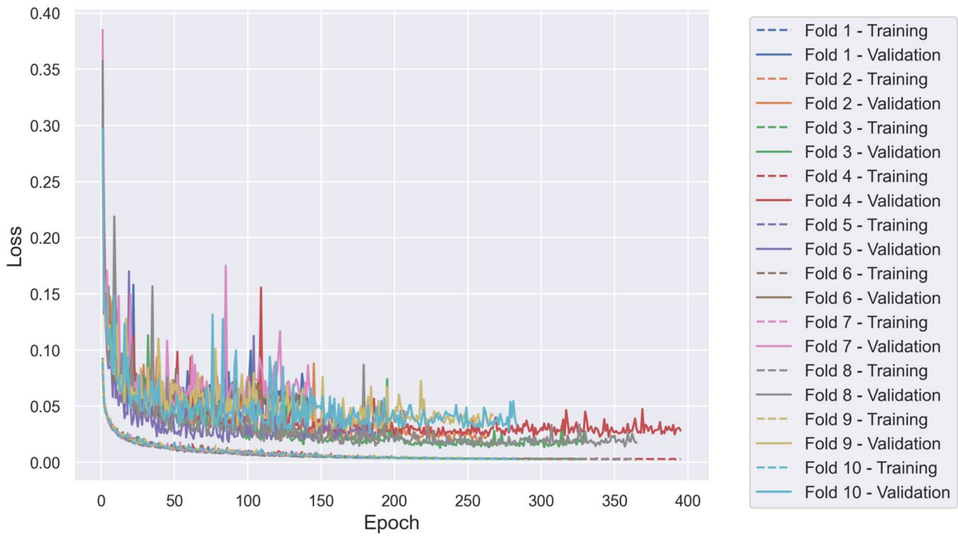

Figure 9. Loss vs. epoch plot of the proposed lightweight CNN with a bottleneck block on the UCI-HAR data set

Table 2. Performance metrics on the UCI-HAR data set sorted using the F1 score of testing

Model	F1 Score		Accuracy		Recall		Precision	
	Testing	**Validation**	**Testing**	**Validation**	**Testing**	**Validation**	**Testing**	**Validation**
Proposed lightweight CNN with convolutional block	0.9625	0.9927	0.9627	0.9927	0.9627	0.9927	0.9641	0.9927
Original lightweight CNN (Ronald et al., 2021)	0.9623	0.9912	0.9624	0.9912	0.9624	0.9912	0.9637	0.9912
Proposed lightweight CNN with identity block	0.9609	0.9927	0.9612	0.9927	0.9612	0.9927	0.9625	0.9927
Proposed lightweight CNN with bottleneck block	0.9557	0.9924	0.9559	0.9924	0.9559	0.9924	0.9577	0.9924

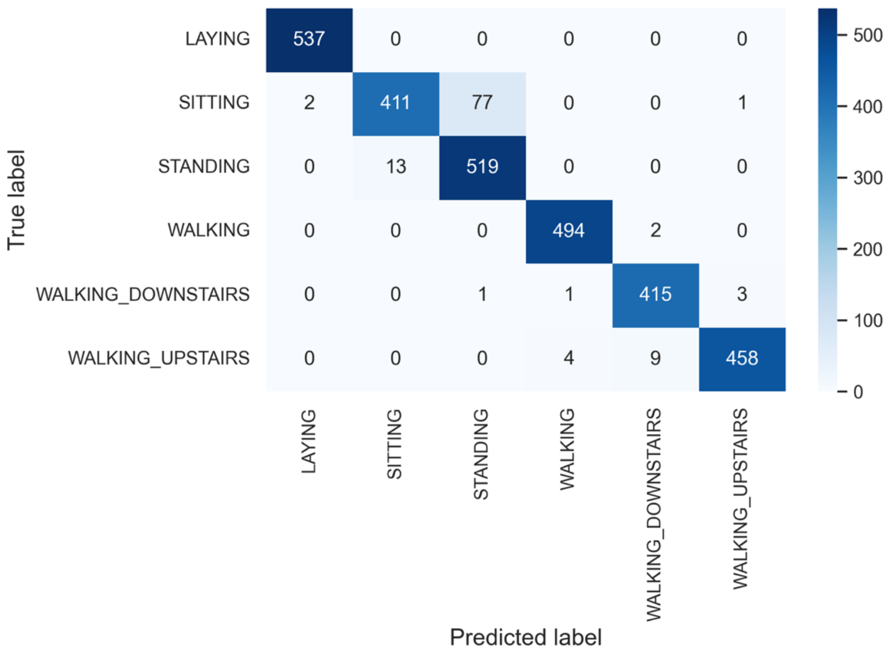

Figure 10. The best confusion matrix for fold 8 of the proposed lightweight CNN with a convolutional block on the UCI-HAR data set

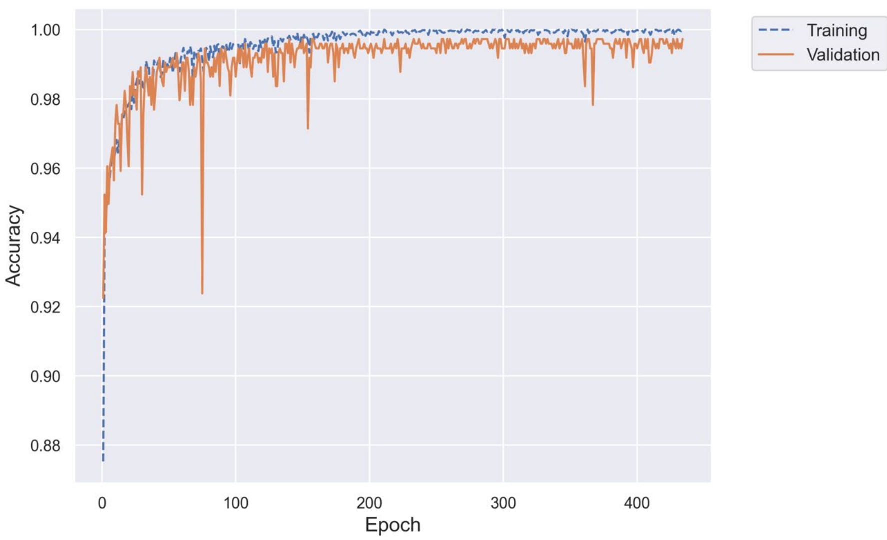

Figure 11. Accuracy vs. epoch plot of fold 8 of the proposed lightweight CNN with a convolutional block on the UCI-HAR data set

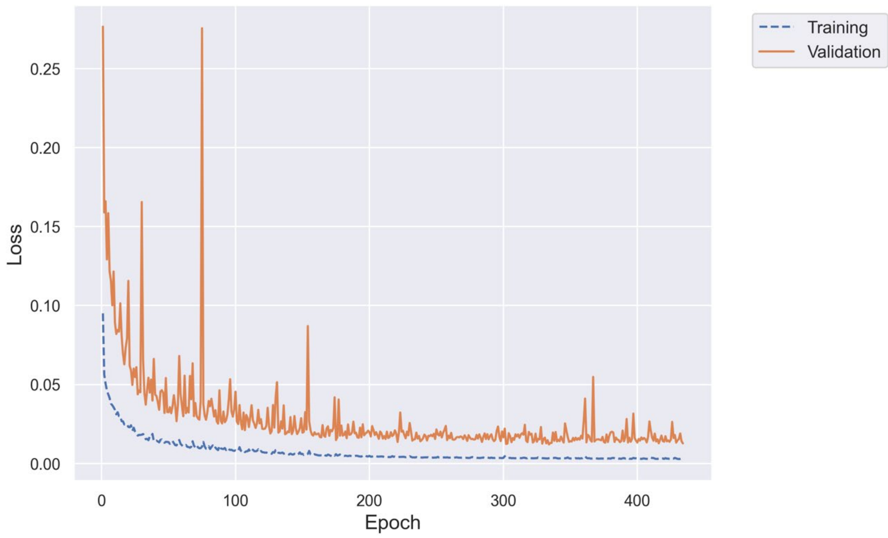

Figure 12. Loss vs. epoch plot of fold 8 of the proposed lightweight CNN with a convolutional block on the UCI-HAR data set

3.2 WISDM

The evaluation of the models on the WISDM data set offers further insight into their performance. Figures 13–20 show the evolution of the accuracy and loss across epochs for all four models, including the original lightweight CNN and the three proposed lightweight CNNs with integrated blocks. Table 3 lists the average performance metrics for each model for the WISDM data set. Interestingly, the proposed lightweight CNN with an identity block delivered the highest F1 scores of 0.9520 and 0.9553 during testing and validation, respectively. This finding was attributed to the proposed lightweight CNN with a convolutional block, with F1 scores of 0.9511 (testing) and 0.9544 (validation). The proposed lightweight CNN with a bottleneck block and the original lightweight CNN exhibited slightly lower performance, achieving F1 scores of 0.9471 and 0.9445 for testing and 0.9513 and 0.9521 for validation, respectively. This pattern was mirrored across the accuracy, recall, and precision metrics, implying that the identity block model slightly outperformed the others. Figure 21 shows the optimal confusion matrix for fold 3 of the proposed lightweight CNN with an identity block on the WISDM data set. In addition, Figures 22 and 23 show the accuracy vs. epoch and loss vs. epoch plots, respectively, for the same fold and model on the WISDM data set.

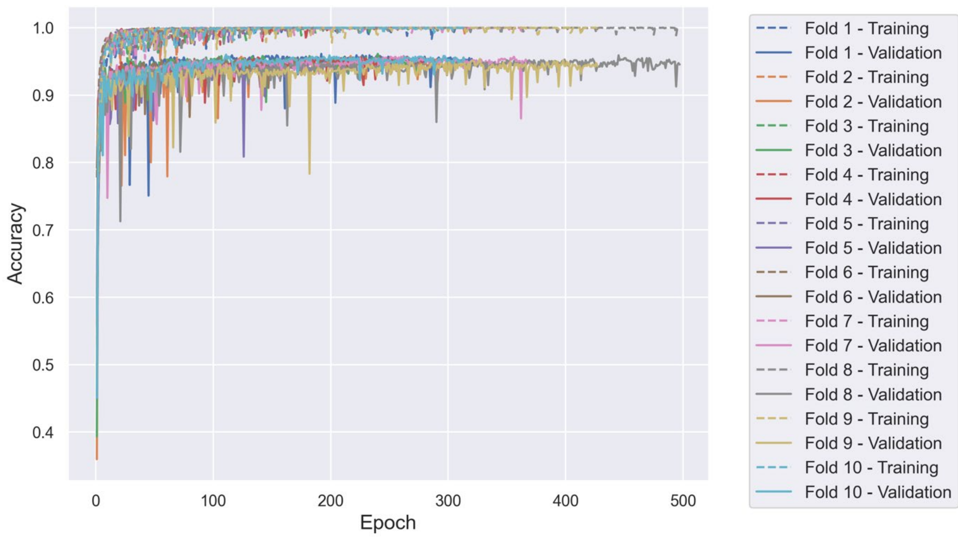

Figure 13. Accuracy vs. epoch plot of the original lightweight CNN on the WISDM data set

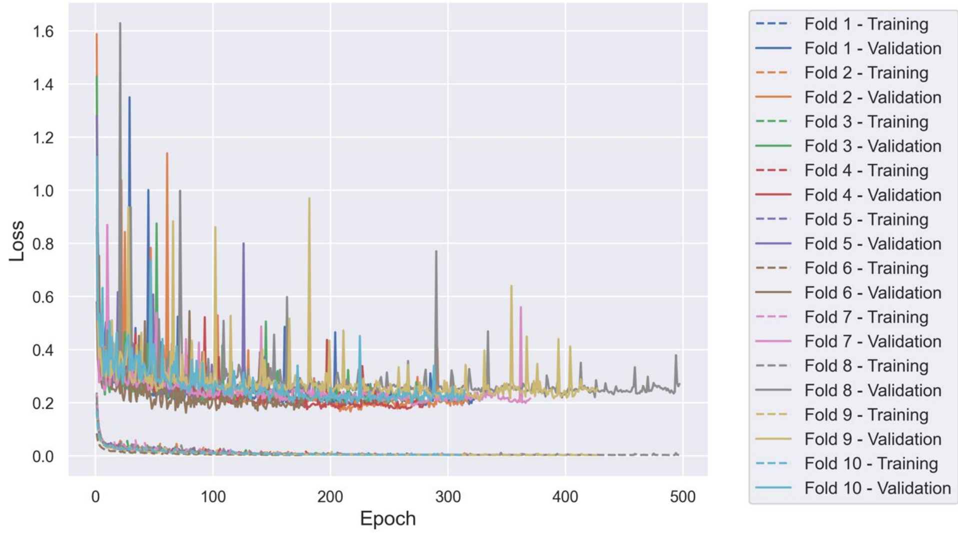

Figure 14. Loss vs. epoch plot of the original lightweight CNN on the WISDM data set

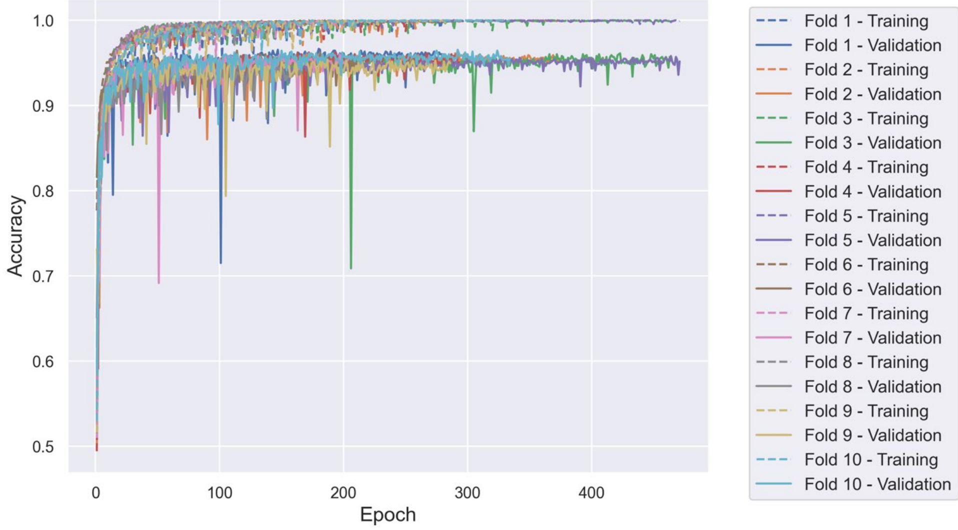

Figure 15. Accuracy vs. epoch plot of the proposed lightweight CNN with an identity block on the WISDM data set

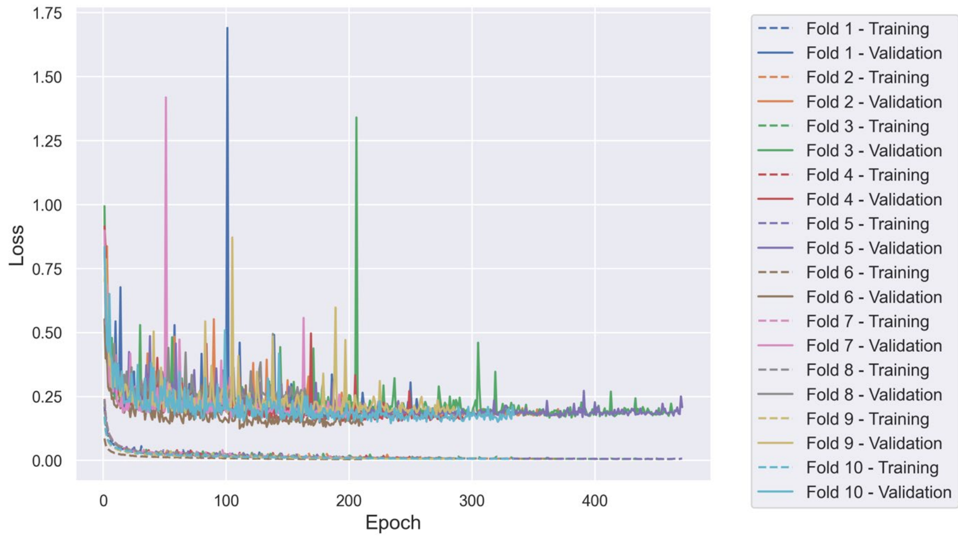

Figure 16. Loss vs. epoch plot of the proposed lightweight CNN with an identity block on the WISDM data set

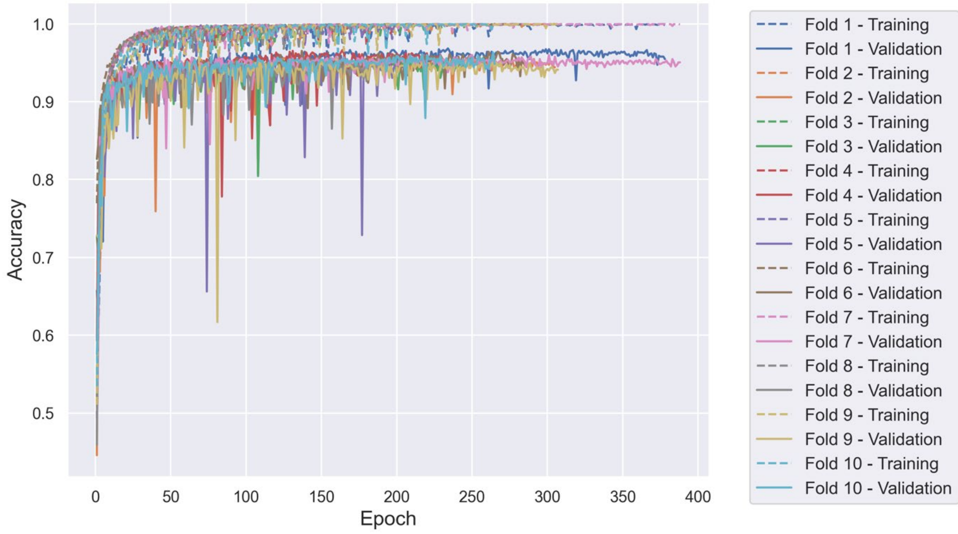

Figure 17. Accuracy vs. epoch plot of the proposed lightweight CNN with a convolutional block on the WISDM data set

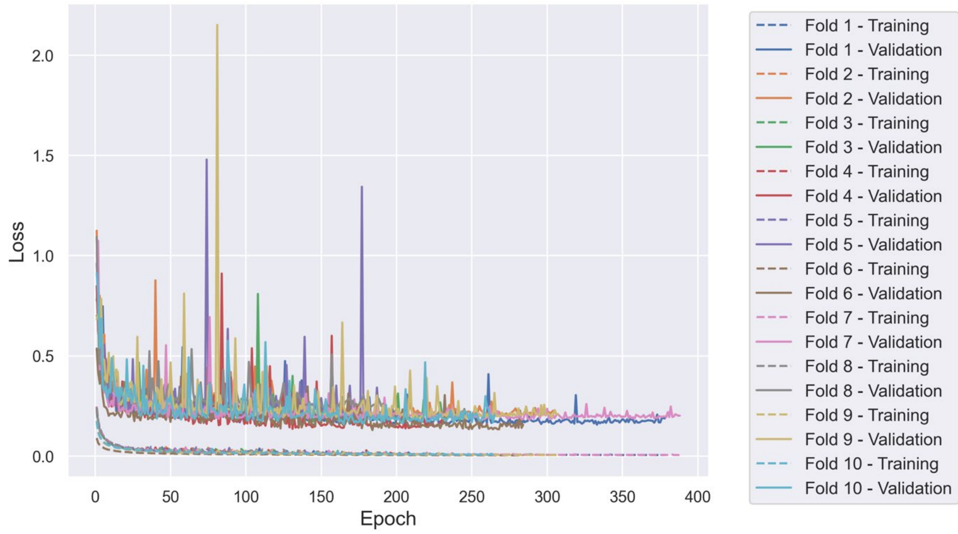

Figure 18. Loss vs. epoch plot of the proposed lightweight CNN with a convolutional block on the WISDM data set

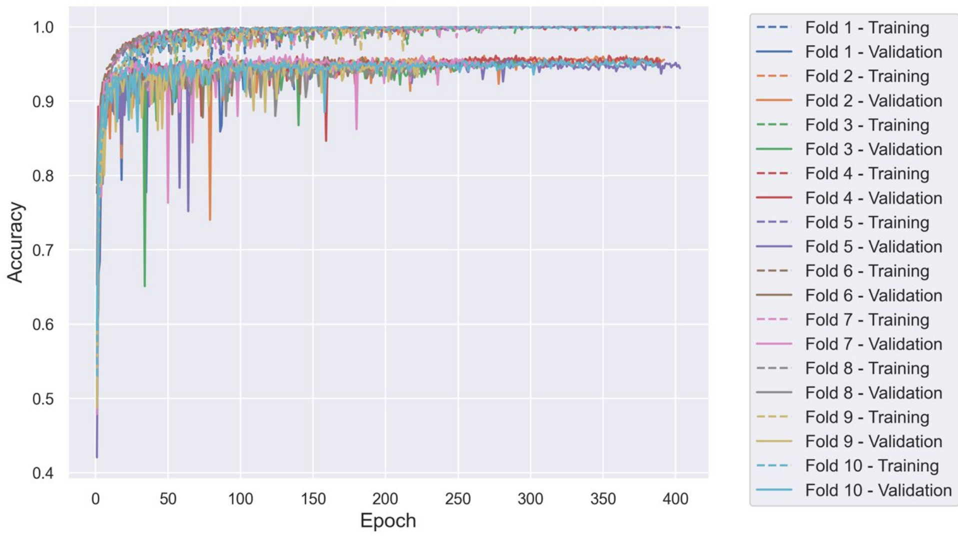

Figure 19. Accuracy vs. epoch plot of the proposed lightweight CNN with a bottleneck block on the WISDM data set

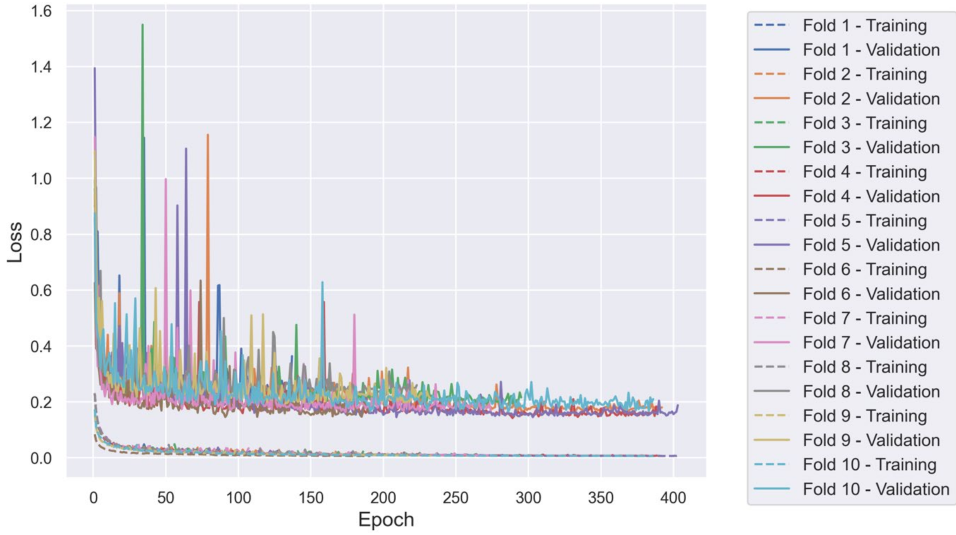

Figure 20. Loss vs. epoch plot of the proposed lightweight CNN with a bottleneck block on the WISDM data set

Table 3. Performance metrics on the WISDM data set sorted by the F1 score of testing

Model	F1 Score		Accuracy		Recall		Precision	
	Testing	Validation	Testing	Validation	Testing	Validation	Testing	Validation
Proposed lightweight CNN with identity block	0.9520	0.9553	0.9523	0.9554	0.9520	0.9556	0.9523	0.9554
Proposed lightweight CNN with convolutional block	0.9511	0.9544	0.9513	0.9546	0.9511	0.9549	0.9513	0.9546
Proposed lightweight CNN with bottleneck block	0.9471	0.9513	0.9472	0.9513	0.9477	0.9520	0.9472	0.9513
Original lightweight CNN (Ronald et al., 2021)	0.9445	0.9521	0.9450	0.9522	0.9448	0.9523	0.9450	0.9522

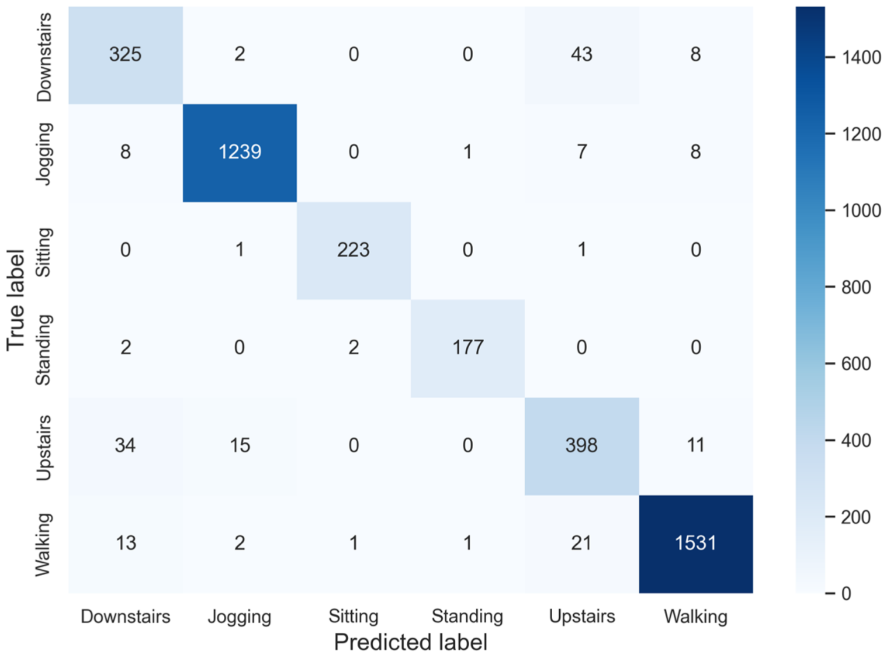

Figure 21. The best confusion matrix for fold 3 of the proposed lightweight CNN with an identity block on the WISDM data set

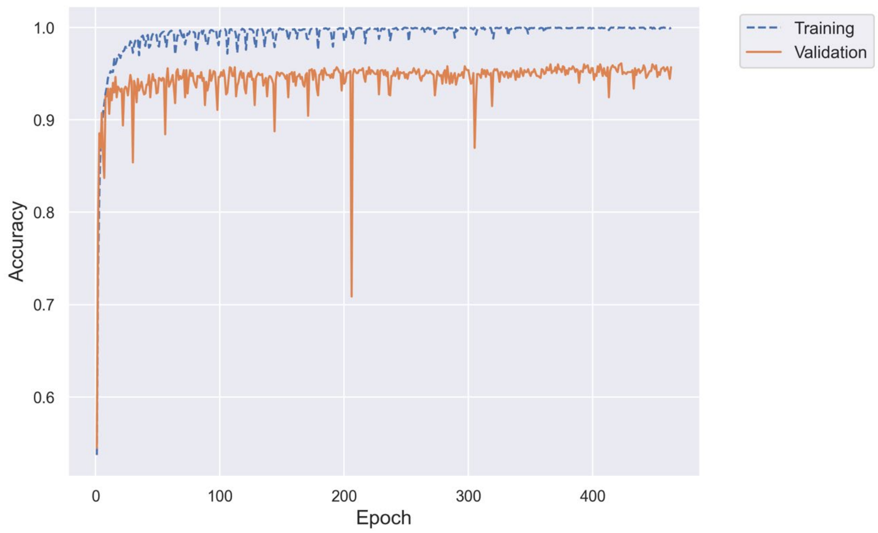

Figure 22. Accuracy vs. epoch plot of fold 3 of the proposed lightweight CNN with an identity block on the WISDM data set

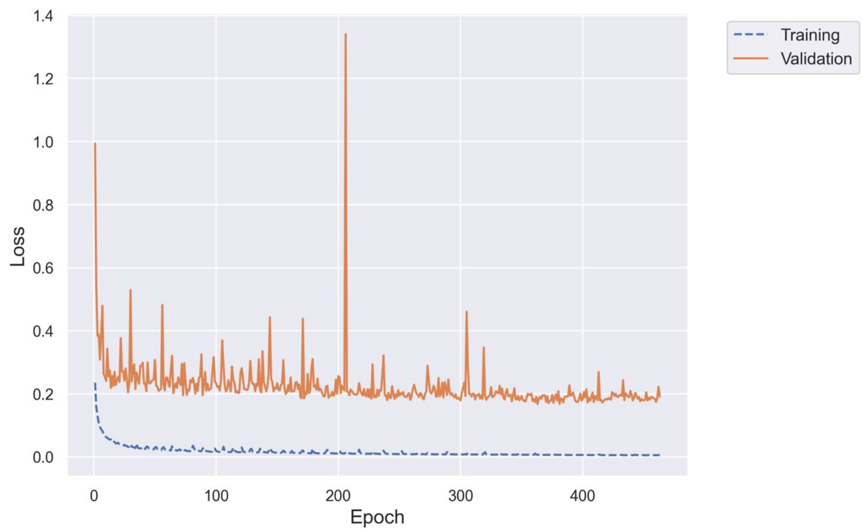

Figure 23. Loss vs. epoch plot of fold 3 of the proposed lightweight CNN with an identity block on the WISDM data set

4. DISCUSSION

The analysis of the proposed models provides valuable insights. The lightweight CNN with integrated blocks, particularly the convolutional and identity block models, outperformed the original lightweight CNN model (Ronald et al., 2021) across both the UCI-HAR and WISDM data sets. The accuracy and epoch plots exhibited fluctuations, reflecting the dynamic interplay between model training and data variability, arising from the inherent variability introduced by K-fold validation to train and evaluate different data subsets.

On the UCI-HAR data set, the convolutional block model recorded the highest F1 score, closely followed by the original lightweight CNN and the identity block models, which suggested minor performance differences among them. Despite exhibiting the least impressive results, the bottleneck block model demonstrated a high F1 score, highlighting its robustness (Wang et al., 2023).

The identity block model marginally surpassed the convolutional block model upon examining the WISDM data set, which was superior to the UCI-HAR data set. These results illustrate that model performance can vary with the data set, suggesting that the optimal model choice can vary based on specific data set characteristics (Yin et al., 2022). The bottleneck block model and original lightweight CNN exhibited comparatively lower performance.

A significant observation is the consistent performance patterns exhibited by all models across the four metrics, suggesting a uniform performance regardless of the evaluation metric used. Another important factor was that the model validation scores consistently surpassed the testing scores. This finding indicates that the models possess good generalization capabilities and can adapt aptly to unseen data. Evidently, even minor differences between the models can potentially become significant in specific applications where slight improvements may lead to a substantial impact (Wang et al., 2023).

The results revealed that the proposed lightweight CNN with a convolutional block outperformed other models, considering the F1 score for the UCI-HAR data set, whereas the proposed lightweight CNN with an identity block achieved the highest F1 score on the WISDM data set. Despite these slight improvements, they are significant when considering the computational complexity. As indicated in Figure 1, the lightweight CNN models with integrated blocks exhibited a strong performance while simultaneously reducing the computational requirements, making them a suitable choice for deployment in devices with limited resources (Chen and Shen, 2017).

However, the most suitable model may vary based on the data set, underlining the significance of tailoring optimizations to specific data sets (Raziani and Azimbagirad, 2022). Despite minor variations in performance, all models exhibited commendable and consistent results across different evaluation metrics. Therefore, the lightweight CNN models with integrated blocks hold substantial potential for future research and real-world applications in HAR.

In Table 4, the performances of different models are compared, and the proposed models surpass the accuracy of previous state-of-the-art models (Ignatov, 2018; Peppas et al., 2020; Ronald et al., 2021). This finding further strengthens the conclusion that the proposed lightweight CNN models with integrated blocks offer viable alternatives to the existing models for HAR tasks.

Table 4. Comparison of model performance

Model	Accuracy	
	UCI-HAR	WISDM
CNN (Ignatov, 2018)	0.9531	0.9332
CNN (Peppas et al., 2020)	-	0.9436
iSPL Inception (Ronald et al., 2021)	0.9509	-
Original CNN (Ronald et al., 2021)	0.9624	0.9450
Proposed lightweight CNN with convolutional block	**0.9627**	0.9513
Proposed lightweight CNN with identity block	0.9612	**0.9523**

5. CONCLUSION

This investigation into lightweight CNN models with integrated blocks demonstrated their potential for improved HAR. Notably, the convolutional and identity block models outperformed the original lightweight CNN model across the UCI-HAR and WISDM data sets. However, the optimal model may differ, depending on the data set. All models, despite minor variations, highlight robust performance across various evaluation metrics, balancing high performance with reduced computational complexity, rendering them appropriate for environments with limited computing resources.

Although the findings of this study are promising, they also recognize certain limitations and suggest directions for future research. One such limitation is the focus on only two data sets, emphasizing the necessity for broader testing across diverse data sets to ensure the generalizability of the proposed models. The second step involves the optimization of the integrated blocks to further enhance the overall model performance and efficiency. In addition, exploring real-world implementations could provide an intriguing avenue for future investigation.

ACKNOWLEDGMENT

The authors would like to acknowledge the providers of the UCI-HAR and WISDM public data sets. The UCI-HAR data set has been ethically approved by UC Irvine, and the WISDM data set has been ethically approved by the Wireless Sensor and Data Mining lab.

REFERENCES

Agac, S., and Incel, O. D. (2023). On the use of a convolutional block attention module in deep learning-based human activity recognition with motion sensors. *Diagnostics (Basel)*, 13(11), 1861.

Barakbayeva, T., and Demirci, F. M. (2023). Fully automatic CNN design with inception and ResNet blocks. *Neural Computing and Applications*, 35(2), 1569–1580.

Chen, Y., and Shen, C. (2017). Performance analysis of smartphone-sensor behavior for human activity recognition. *IEEE Access*, 5, 3095–3110.

Gupta, N., Gupta, S. K., Pathak, R. K., Jain, V., Rashidi, P., and Suri, J. S. (2022). Human activity recognition in artificial intelligence framework: A narrative review. *Artificial Intelligence Review*, 55, 4755–4808.

Ignatov, A. (2018). Real-time human activity recognition from accelerometer data using convolutional neural networks. *Applied Soft Computing*, 62, 915–922.

Ismail, W. N., Alsalamah, H. A., Hassan, M. M., and Mohamed, E. (2023). AUTO-HAR: An adaptive human activity recognition framework using an automated CNN architecture design. *Heliyon*, 9(2), e13636.

Kashyap, S. K., Mahalle, P. N., and Shinde, G. R. (2022). Human activity recognition using 1-Dimensional CNN and comparison with LSTM. *Lecture Notes in Electrical Engineering*, 939, 1017–1030.

Murad, A., and Pyun, J.-Y. (2017). Deep recurrent neural networks for human activity recognition. *Sensors*, 17(11), 2556.

Negi, A., Kumar, K., Chaudhari, N. S., Singh, N., and Chauhan, P. (2021). Predictive analytics for recognizing human activities using residual network and fine-tuning. In *Big Data Analytics. BDA 2021. Lecture Notes in Computer Science* (Srirama, S. N., Lin, J. C. W., Bhatnagar, R., Agarwal, S., and Reddy, P. K., Eds.), pp. 296–310. Berlin: Springer.

Peppas, K., Tsolakis, A. C., Krinidis, S., and Tzovaras, D. (2020). Real-time physical activity recognition on smart mobile devices using convolutional neural networks. *Applied Sciences*, 10(23), 8482.

Phukan, N., Mohine, S., Mondal, A., Manikandan, M. S., and Pachori, R. B. (2022). Convolutional neural network-based human activity recognition for edge fitness and context-aware health monitoring devices. *IEEE Sensors Journal*, 22(22), 21816–21826.

Raziani, S., and Azimbagirad, M. (2022). Deep CNN hyperparameter optimization algorithms for sensor-based human activity recognition. *Neuroscience Informatics*, 2(3), 100078.

Reyes-Ortiz, J., Anguita, D., Ghio, A., Oneto, L., and Parra, X. (2012). *Human Activity Recognition Using Smartphones*. [Online URL: https://archive.ics.uci.edu/dataset/240/human+activity+recognition+using+smartphones] accessed on September 25, 2023.

Ronald, M., Poulose, A., and Han, D. S. (2021). iSPLInception: An inception-resnet deep learning architecture for human activity recognition. *IEEE Access*, 9, 68985–69001.

Souza, R. M., Nascimento, E. G. S., Miranda, U. A., Silva, W. J. D., and Lepikson, H. A. (2021). Deep learning for diagnosis and classification of faults in industrial rotating machinery. *Computers and Industrial Engineering*, 153, 107060.

Straczkiewicz, M., James, P., and Onnela, J.-P. (2021). A systematic review of smartphone-based human activity recognition methods for health research. *npj Digital Medicine*, 4(1), 1–15.

Teng, Q., Zhang, L., Tang, Y., Song, S., Wang, X., and He, J. (2021). Block-wise training residual networks on multi-channel time series for human activity recognition. *IEEE Sensors Journal*, 21, 18063–18074.

Wang, A., Chen, G., Yang, J., Zhao, S., and Chang, C.-Y. (2016). A comparative study on human activity recognition using inertial sensors in a smartphone. *IEEE Sensors Journal*, 16(11), 4566–4578.

Wang, J., Chen, Y., Hao, S., Peng, X., and Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119, 3–11.

Wang, Y., Xu, H., Zheng, L., Zhao, G., Liu, Z., Zhou, S., Wang, M., and Xu, J. (2023). A multi-dimensional parallel convolutional connected network based on multi-source and multi-modal sensor data for human activity recognition. *IEEE Internet of Things Journal*. 10(16), 14873–14885.

Weiss, G. (2019). *WISDM Smartphone and Smartwatch Activity and Biometrics Dataset*. [Online URL: https://archive.ics.uci.edu/dataset/507/wisdm+smartphone+and+smartwatch+activity+and+biometrics+dataset] accessed on September 25, 2023.

Xu, Z., Zhao, J., Yu, Y., and Zeng, H. (2020). Improved 1D-CNNs for behavior recognition using wearable sensor network. *Computer Communications*, 151, 165–171.

Yin, X., Liu, Z., Liu, D., and Ren, X. (2022). A novel CNN-based Bi-LSTM parallel model with attention mechanism for human activity recognition with noisy data. *Scientific Reports*, 12(1), 1–11.

Zhongkai, Z., Kobayashi, S., Kondo, K., Hasegawa, T., and Koshino, M. (2022). A comparative study: Toward an effective convolutional neural network architecture for sensor-based human activity recognition. *IEEE Access*, 10, 20547–20558.

Zhou, Y., Chen, S., Wang, Y., and Huan, W. (2020). Review of research on lightweight convolutional neural networks. In *Proceedings of 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 1713–1720. Chongqing, China.

Zhou, Y., Wang, X., Zhang, M., Zhu, J., Zheng, R., and Wu, Q. (2019). MPCE: A maximum probability based cross entropy loss function for neural network classification. *IEEE Access*, 7, 146331–146341.