Federated Learning and Genetic Mutation for Multi-Resident Activity Recognition

Cezar Anicai, Muhammad Zeeshan Shakir
cezar.anicai@uws.ac.uk, muhammad.shakir@uws.ac.uk
School of Computing, Engineering and Physical Sciences
University of the West of Scotland
Paisley, United Kingdom

***Abstract*—Multi-Resident activity recognition refers to the task of recognizing activities performed by multiple individuals living in the same residence. It involves using sensors or other monitoring devices to capture data about the activities taking place in the living space, and then using Machine Learning (ML) or Deep Learning (DL) algorithms to analyze and classify these activities. Federated Learning (FL) is a technique that enables multiple devices to collaboratively train a model without sharing their data with each other, while Genetic Mutation (GM) is a technique used in evolutionary algorithms to introduce random changes to the genetic code of individuals in a population. Our proposed framework involves the use FL and GM for Human Activity Recognition (HAR). The approach was evaluated on the ARAS dataset, collected from two houses with different activity patterns. Two Recurrent Neural Network (RNN) models, Gated Recurrent Unit (GRU) and Long-Short Term Memory (LSTM), were employed for the activity classification task and a genetic mutation operator was applied to the weights of the models before federated averaging. The results indicate that FL is suitable for privacy preserving activity recognition, it can help with early deployment and even improve the performance of the models in some cases.**

***Index Terms*—activity recognition, multi-resident, deep learning, federated learning, genetic mutation, LSTM, GRU, ARAS.**

I. Introduction

The field of Ambient Assisted Living (AAL) has gained popularity in the recent years due to the benefits that AAL systems could bring to quality of life. One major direction of research and development within this field is Human Activity Recognition (HAR), which attempts to automatically detect various activities performed by humans in real life scenarios, by employing intelligent analysis on data from various sources and of varied types [1].

Internet of Things (IoT) is one of the enabling technologies for activity recognition because it provides the means for acquiring and transmitting data. For this task in particular, video and motion capturing equipment such as cameras, motion, distance or force sensors are tipically deployed [2]. Wireless communication protocols like Bluetooth, Wi-Fi, SigFox and others are then used to transmit collected data to a central processing platform, often identified as cloud servers.

Deep Learning (DL) is another enabling technology for activity recognition, from an intelligent analysis perspective. DL is a subfield of machine learning that covers any type of Neural Network (NN) model, including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Generative Adversarial Network (GAN) or Autoencoders. It has been proven to be particularly effective for human activity recognition tasks due to the ability of DL models to extract complex features from raw data, such as accelerometer or image data, without requiring manual feature engineering [3], [4]. DL models are able to automatically learn hierarchical representations of the data, which can capture both low-level and high-level features. This allows them to perform well on tasks where the relationships between the input data and the output labels are complex and non-linear, such as recognizing different human activities from sensor data.

DL models for HAR typically require a large amount of labeled data, which can be challenging to obtain, especially in the case of multi-resident homes. Federated Learning (FL) enables algorithm training using data that is distributed across multiple devices, without the need for centralized data storage, and has emerged as a promising solution to this challenge.

In this paper, we propose the use of FL for multi-resident activity recognition on the Activity Recognition with Ambient Sensing (ARAS) dataset [5], which consists of sensor data collected from a smart home environment. We develop a FL algorithm that allows for the collaborative training of two deep neural networks, LSTM and GRU, using data from multiple residents in a privacy-preserving manner.

Our proposed approach is evaluated on the ARAS dataset and the experimental results show that FL can bring a slight improvement in performance especially when coupled with a genetic mutation operator. The time cost of our method is in the order of milliseconds but depending on the use case, having a privacy preserving intelligent analysis can be worth the delays introduced by the additional operations required. Moreover, if we take in consideration the time and bandwidth required for aggregating training data to a central server, our approach could prove more efficient overall.

The remainder of this paper is organized as follows. In section II, we provide a brief overview of related work in the domain of activity recognition and FL. In section III, we describe the proposed approach in detail, including the architecture of the DL algorithms, the FL algorithm and the mutation operator. In section IV, we present the experimental results in terms of accuracy and execution time. Finally, we conclude the paper in section V with a discussion of the main contributions and future directions for this research topic.

II. Related Works

A. Deep Learning in Activity Recognition

The authors in [6] proposed a hybrid approach for HAR by employing a Support Vector Machine (SVM) algorithm for static activity recognition and a CNN model for moving activity recognition. The initial split of activities in static and dynamic/moving is done with a Random Forest (RF) binary classifier, the CNN model has three convolution layers, Rectified Linear Unit (ReLU) activation and 20% dropout rate and the SVM is a one-vs-one three class classification model. The approach was evaluated on the UCI-HAR dataset and obtained an accuracy score of 97.66%. A study on Deep Convolutional Neural Network (DCNN) applied to micro-Doppler signature images is presented in [7]. Three pre-trained DCNN models, VGG-16, VGG-19 and Inception V3 were fine-tuned and tested on a dataset consisting of 6 activities. The accuracies obtained by the three models were reported as 99.3%, 95.8%, and 97.38% respectively, and among the hyperparameters chosen during fine-tuning we can find a learning rate of 0.0001, Leaky ReLU as activation function and RMSProp as optimizer. The authors in [8] used the accelerometer and gyroscope sensors embedded in mobile devices to collect data about Activities of Daily Living (ADLs) from 12 individuals. The resulting dataset included 94198 observations and six activities, sitting, running, standing, walking, going upstairs and going downstairs. A hybrid CNN-LSTM model was implemented to efficiently handle hierarchical features, the activation function used for both CNN and LSTM layers was ReLU and the loss function used in training was categorical cross-entropy. The performance of the model on their own dataset was reported as 97.61% accuracy while testing on a second public dataset, also with 6 ADLs, reported an average accuracy of 98.69%. A comparison between Unidirectional and Bidirectional LSTM models on two different accelerometer datasets is provided in [9]. The results showed that the bidirectional approach has a higher recognition power with an average accuracy of 92.63% over the two datasets compared to 91.73% for the unidirectional approach.

B. Federated Learning in Activity Recognition

The authors in [10] proposed a Graph Convolution Network trained in federated settings to combat issues like privacy preservation and label scarcity in HAR tasks. The Graph Convolution Network uses inter-relatedness and closeness of activities to propagated activity labels within the datasets in a semi-supervised manner. Testing the method on two datasets revealed an average accuracy of 89.8% for the FL approach while traditional centralized training managed an accuracy of 94.85%. The study from [11] compared the performance and communication cost of FL to centralized learning for HAR on synthetic and on real-world datasets. They built two models, a Deep Neural Network (DNN) and a softmax regression, the former achieved an accuracy of 89% in federated settings and 93% in centralized training while the later managed an accuracy of 80% in federated settings compared to 83% in centralized training. Additionally, the communication cost for FL has been shown to increase with the complexity of the model. A federated clustering technique is described in [12] to tackle the issue of non-Independently and Identically Distributed (non-IID) data in HAR. The proposed algorithm, named FedCLAR, groups local clients based on their similarity and allows clients in the same cluster to generate a refined shared model for their cluster. The remaining clients use a global model generated by averaging the weights of all local models regardless of their belonging to a cluster. Testing the approach on two public datasets revealed that it outperforms other state-of-the-art FL methods. The authors in [13] proposed P-FedAvg algorithm to tackle the global model bias caused by direct averaging in FL and to improve the accuracy of local models aggregation. P-FedAvg uses an additional parameter P during the aggregation process to adjust the weights of participating models based on their sample size. Experiments on the UCI-HAR dataset showed their approach improves the accuracy of FL models in comparison to simple federated averaging. Study [14] presented a comparison between Federated Averaging, Federated Personalization and Federated Match Averaging. The testing done on HAR datasets showed Federated Averaging as the most effective technique.

C. Multi-Resident Activity Recognition

Arguing that in real-world scenarios the activities that need to be recognized can be previously unseen activities by the trained models, [15] proposed a method involving multi-task learning and zero-shot learning for multi-resident HAR. Experiments on the publicly available ARAS dataset demonstrated the effectiveness of the approach. The authors in [16] implemented a Classifier Chain with K-Nearest Neighbor algorithm to classify multi-resident activities in consideration to underlying dependencies between the activities of different resident from the same facility. Testing the approach on the above mentioned ARAS dataset, showed accuracy levels between 75.8% and 93.1%. The study in [17] did a comparative analysis of Feature Selection (FS) techniques in the case of multi-resident HAR. They considered three FS approaches, filter-based (Information Gain), wrapper–based (Recursive Feature Elimination) and embedded–based (Tree-based). Each of these methods was combined with five Machine Learning (ML) algorithms, Logistic Regression, Linear Discriminant Analysis, Naïve Bayes, RF, and K-Nearest Neighbor. Evaluation on a publicly available dataset proved that RF combined with Tree-based FS outperforms all of the other considered combinations. A RGB image-based DCNN classifier for unobtrusive activity recognition was proposed by [18]. The stucture of the network consisted in three convolutional layers followed by a max pooling layer for feature extraction and three fully connected layers for classification. The model was evaluated on the Cairo dataset provided by the CASAS project, a dataset with 13 activities collected from two residents. The results showed this particular DCNN achieves a peak accuracy of 95.2%, higher than previous work on the same dataset.

III. Proposed Framework

A. Dataset and data preprocessing

We considered for our framework the ARAS dataset introduced by [5] in 2013. The data was collected from two houses with two residents each for a period of one month. Both houses were equipped with 20 binary sensors of the following types, force sensitive resistors, pressure mats, contact sensors, proximity sensors, sonar distance sensors, photocells, temperature sensors and infrared receivers. The sensors recorded one sample every second bringing the total number of observations for each house to 2,592,000. A number of 27 different activities were recorded and they are listed in Table I. For the reminder of this work we will refer the two houses as House A and House B. It is important to note that data was collected from two different households with different activity patterns, a closer look at the dataset revealed that activity 26, having guests, was not present at all in House B so we also removed it from House A in order to keep the features relevant for FL. Further more, the order of the features from the two houses was not consistent and we had to reorder the columns to match their counterpart from the other house in order to avoid meaningless averaging of the local models. In terms of data cleanliness and quality, no missing values or null observations were found and no normalization was needed since the features are binary in value.

B. Deep Learning

In terms of intelligent analysis we choose two RNN models, GRU and LSTM, due to their previously proven good performance on HAR tasks and their suitability for federated averaging. A basic RNN cell structure is depicted in Fig. 1, where the input x_t and previous hidden state h_{t-1} are combined into a vector. The values of the vector pass through a tanh activation function outputting the h_t hidden state of the current cell.

LSTM has a more complex cell structure which allows the network to keep or forget previous information, see Fig. 2. LSTM introduces a cell state, noted with c_t/c_{t-1}, which acts as the memory of the network and can carry information from the earlier time steps to much later time steps, hence the "long" memory name. The amount of information carried forward is decided by a forget gate f_t, which applies a sigmoid activation function to the previous hidden state and the current input. The output of the sigmoid function varies between 0 and 1, a value closer to 0 means forget and a value closer to 1 means remember. An input gate i_t is used to update the value of the cell state through the same mechanism used in the forget gate, along with a candidate cell state c_t obtained by passing the previous hidden state and the input through a tanh activation function. Finally, an output gate o_t calculates the new hidden state h_t, taking into account the previous hidden state, the cell state and the input.

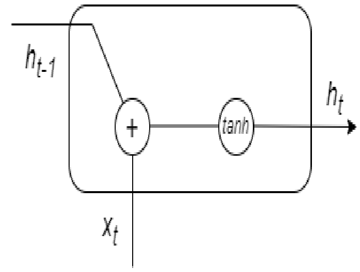

Fig. 1. RNN cell

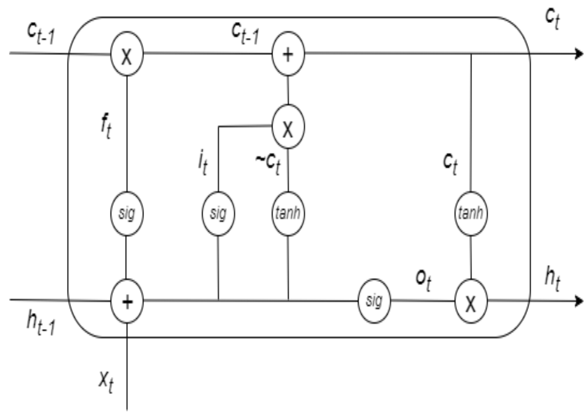

Fig. 2. LSTM cell

TABLE I
ACTIVITY TYPES

ID	Activity	ID	Activity	ID	Activity
1	Other	10	Having Snack	19	Laundry
2	Going Out	11	Sleeping	20	Shaving
3	Preparing Breakfast	12	Watching TV	21	Brushing Teeth
4	Having Breakfast	13	Studying	22	Talking on the Phone
5	Preparing Lunch	14	Having Shower	23	Listening to Music
6	Having Lunch	15	Toileting	24	Cleaning
7	Preparing Dinner	16	Napping	25	Having Conversation
8	Having Dinner	17	Using Internet	26	Having Guest
9	Washing Dishes	18	Reading Book	27	Changing Clothes

GRU is a newer type of RNN similar to LSTM but arguably faster and more lightweight in terms of computation. GRU eliminated the need for a cell state and rather uses the hidden state to carry on previous information. It introduced an update gate u_t which works similar to the input gate from LSTM and a reset gate r_t which decides how much information to forget or carry on. A representation of its cell structure can be observed in Fig. 3.

C. Federated Learning and Genetic Mutation

Genetic algorithms are a class of optimization algorithms that are inspired by the process of natural selection in biology.

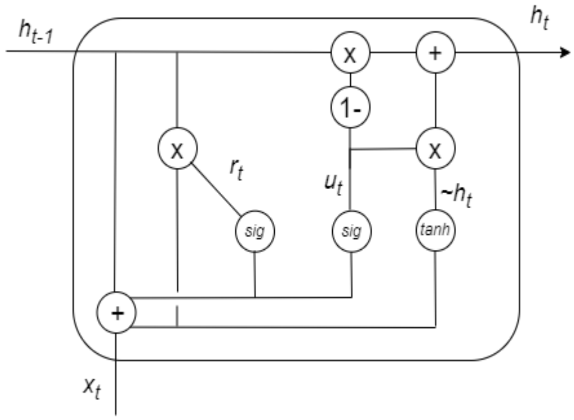

Fig. 3. GRU cell

They are used to find solutions to optimization problems by generating a population of candidate solutions, and then iteratively selecting the best ones and recombining them to create new candidate solutions. One of the key operations in genetic algorithms is Genetic Mutation (GM). GM is the process by which the genetic material of an individual in the population is randomly altered to create a new solution. In genetic algorithms, mutation is used to introduce diversity into the population and to explore the search space more thoroughly. Hence, we choose to apply a mutation operator to our models by altering the local model weights after they have trained but before federated averaging generates a global model. The mutation operator is described in Eq. 1 and Eq. 2, where w_{max}, w_{min} are the maximum and minimum values of the weights within one network layer, w is the current value of the weight being mutated, w' is the new value after mutation and $\mathcal{N}(0, \sigma)$ is a random number drawn from a normal distribution with mean 0 and standard deviation σ. The probability of a weight to be mutated was chosen as 5%.

$$\sigma = \frac{w_{max} - w_{min}}{6} \tag{1}$$

$$w' = w + \mathcal{N}(0, \sigma) \tag{2}$$

For the federated averaging process, we first train the algorithms independently on each house's subset of data and perform testing on 10% of each subset. The weights from each layer for both models are then averaged and the resulting model is tested against the house with a poorer performance during the local training. A visual explanation for the information flow in our proposed framework is presented in Fig. 4.

IV. Results and Discussion

We split both House A and House B datasets into 6 subsets each, containing 5 days worth of data (432,000 observations), we further split all the subsets in 90% training data and 10% test data and fitted the LSTM and GRU models. Since we

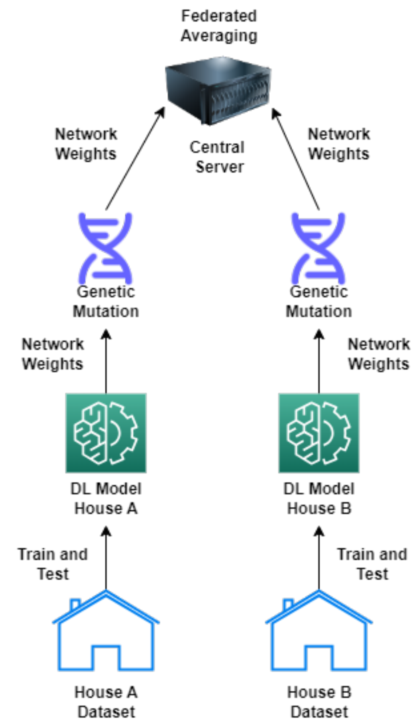

Fig. 4. Federated Learning and Genetic Mutation Diagram

have two houses, we implemented one model for each house, resulting in two LSTM networks and two GRU networks. The structure of the networks was consistent across the two houses to allow direct averaging of the weights, a discrepancy in the number of layers or units within a layer would yield a different number of weights for the models.

For the LSTM models we implemented a structure with a fully connected layer with 20 units and a LSTM hidden layer with 100 units. The loss function chosen was categorical crossentropy, the activation function was sigmoid, dropout and recurrent dropout values were set to 0.1, the optimizer chosen was Adam, the batch size was set to 25000 given the large size of the dataset and the models were run for 15 epochs. The same structure and parameters were used for the GRU models as well, except for the optimizer which was chosen as AdaDelta.

The accuracy scores for our approach are presented in Table II and Table III. The results are consistent across both models which indicates that the approach can be model independent and work well with other NN. It is important to note that the averaged models were only applied to the house with poorer results. Excluding the special case of the first 5 days subset, where the accuracy for House B was tremendously low and applying federated averaging substantially improved the performance, the simple FL approach improved the accuracy with up to 1.6% for some of the test scenarios, while the worst performance was registered as a decrease of 4.6%. When combined with a genetic mutation operator, our approach

TABLE II
GRU ACCURACY SCORES

	$HouseA$	$HouseB$	FL	FL + GM
$5Days$	82.7%	14.2%	82%	82.4%
$10Days$	75%	90.9%	75.7%	77.7%
$15Days$	75.8%	92.9%	77.4%	76.8%
$20Days$	78.2%	84.7%	77.2%	79.2%
$25Days$	75.2%	87.3%	75%	76.1%
$30Days$	81.2%	93.1%	81.2%	80.1%

TABLE III
LSTM ACCURACY SCORES

	$HouseA$	$HouseB$	FL	FL + GM
$5Days$	87.9%	18%	82%	82%
$10Days$	83.9%	94%	79.3%	83.9%
$15Days$	88.6%	93.6%	86.8%	88%
$20Days$	88.4%	92.7%	88.4%	88%
$25Days$	80.9%	89.6%	81.1%	80.9%
$30Days$	89.4%	93.6%	88.4%	89.4%

TABLE IV
EXECUTION TIME (MS) - LSTM

	FL	FL + GM
$5Days$	17.51	19.94
$10Days$	7.49	7.98
$15Days$	7.98	7.98
$20Days$	7.98	7.98
$25Days$	6.98	7.48
$30Days$	7.98	7.98

TABLE V
EXECUTION TIME (MS) - GRU

	FL	FL + GM
$5Days$	19.45	20.45
$10Days$	6.98	6.98
$15Days$	6.98	7.98
$20Days$	6.98	6.98
$25Days$	7.98	7.98
$30Days$	6.98	7.48

showed a 2.7% improvement in accuracy as the best result while the worst performance was a 1.1% decrease.

The experiments were ran on a Windows 11 machine with a 6 cores Intel i5-10400 CPU with a frequency of 2.90GHz, 16Gb RAM with a speed of 2667MHz and Solid-state drive storage device. The execution times for the federated averaging and the combined federated averaging with weights mutation for each of the two models are shown in Table IV and Table V. As seen in the tables, applying FL introduces a delay ranging from 6.98ms to 19.45ms, however the delay introduced by the mutation operator on top of the federated averaging procedure is minimal. It is important to note that the delay introduced by our approach can be compensated for by eliminating the need to transmit the entire dataset to a central location, which would also increase the bandwidth cost.

V. CONCLUSION

In this paper, we proposed a FL and GM approach for multi-resident activity recognition using the ARAS dataset. Our approach enables the collaborative training of NN using data from multiple residents in a privacy-preserving manner.

The experimental results showed that the proposed approach can bring a performance improvement in some cases, while in some scenarios, privacy preservation will come at the cost of a decrease in performance. From our results we can derive a further use for FL except the privacy preservation for data. In the case of activity recognition, averaging the parameters of models from different households can enable a quicker and more robust deployment. A FL model will learn to recognize activities from multiple local datasets, which allows the model to be used even when a particular local dataset does not have enough observations to represent all the activities equally well.

In conclusion, our proposed approach for multi-resident activity recognition using FL and GM represents an important step forward in the development of smart home applications that can provide support and care for the residents. The approach's ability to leverage the distributed data from multiple residents while ensuring data privacy makes it a promising solution for activity recognition using ambient sensors in multi-resident homes.

REFERENCES

[1] J. H. Shah and A. D. Patel. "Ambient assisted living system: The scope of research and development". In: *2018 international conference on electrical, electronics, computers, communication, mechanical and computing (EECCMC), Tamilnadu, India*. 2018.

[2] A. Gupta et al. "A Survey on Human Activity Recognition and Classification". In: *2020 International Conference on Communication and Signal Processing (ICCSP)*. 2020, pp. 0915–0919.

[3] D. Khan and I. W.-H. Ho. "Deep Learning of CSI for Efficient Device-free Human Activity Recognition". In: *2021 IEEE 7th World Forum on Internet of Things (WF-IoT)*. 2021, pp. 19–24.

[4] S. Waheed et al. "An Automated Human Action Recognition and Classification Framework Using Deep Learning". In: *2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. 2023, pp. 1–5.

[5] H. Alemdar et al. "ARAS human activity datasets in multiple homes with multiple residents". In: *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*. 2013, pp. 232–235.

[6] S.M. Hossain et al. "A Hybrid Approach for Human Activity Recognition with Support Vector Machine and 1D Convolutional Neural Network". In: *2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. 2020, pp. 1–5.

[7] A. Waghumbare, U. Singh, and N. Singhal. "DCNN Based Human Activity Recognition Using Micro-Doppler Signatures". In: *2022 IEEE Bombay Section Signature Conference (IBSSC)*. 2022, pp. 1–6.

[8] N.A. Choudhury and B. Soni. "An Efficient CNN-LSTM Approach for Smartphone Sensor-Based Human Activity Recognition System". In: *2022 5th International Conference on Computational Intelligence and Networks (CINE)*. 2022, pp. 01–06.

[9] L. Alawneh et al. "A Comparison of Unidirectional and Bidirectional LSTM Networks for Human Activity Recognition". In: *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. 2020, pp. 1–6.

[10] A. Sarkar, T. Sen, and A.K. Roy. "GraFeHTy: Graph Neural Network using Federated Learning for Human Activity Recognition". In: *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2021, pp. 1124–1129.

[11] K. Sozinov, V. Vlassov, and S. Girdzijauskas. "Human Activity Recognition Using Federated Learning". In: *2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom)*. 2018, pp. 1103–1111.

[12] R. Presotto, G. Civitarese, and C. Bettini. "FedCLAR: Federated Clustering for Personalized Sensor-Based Human Activity Recognition". In: *2022 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. 2022, pp. 227–236.

[13] S. Xu, Y. Liu, and X. He. "Studies on Human Recognition Activities Based on Federated Learning". In: *2022 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI)*. 2022, pp. 372–377.

[14] S. Ek et al. "Evaluation of Federated Learning Aggregation Algorithms: Application to Human Activity Recognition". In: *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*. 2020, pp. 638–643.

[15] W. Wang and C. Miao. "Multi-Resident Activity Recognition with Unseen Classes in Smart Homes". In: *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*. 2018, pp. 780–787.

[16] M. Jethanandani et al. "Multi-Resident Activity Recognition using Multi-Label Classification in Ambient Sensing Smart Homes". In: *2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*. 2019, pp. 1–2.

[17] J.W. Kasubi and M.D. Huchaiah. "A Comparative Study of Feature Selection Methods for Activity Recognition in the Smart Home Environment". In: *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)*. 2022, pp. 890–895.

[18] T.-H. Tan et al. "Multi-Resident Activity Recognition in a Smart Home Using RGB Activity Image and DCNN". In: *IEEE Sensors Journal* 18.23 (2018), pp. 9718–9727.