Semi-Supervised Learning with Sparsely Labelled Multi-Sensor Activity Data for Wellness Monitoring of Infants

Avirup Roy [1], Amit Kumar Bhuyan [1], Hrishikesh Dutta [1] , Mei-Hua Lee [2], and Subir Biswas [1]
[1]Department of Electrical and Computer Engineering, [2]Department of Kinesiology, Michigan State University, East Lansing, MI, USA
royaviru@msu.edu, bhuyanam@msu.edu, duttahr1@msu.edu, mhlee@msu.edu, sbiswas@egr.msu.edu

***Abstract*— This paper presents a novel on-device semi-supervised learning framework for infant activity classification using sparsely labeled multi-sensor data. The framework employs an unsupervised clustering mechanism, followed by cluster labeling with minimally labeled data. It is specifically designed for wellness monitoring applications that involve infant activity detection. The proposed semi-supervised learning (SSL) model can self-train and predict multiple activity classes without requiring pre-labeled data, while maintaining minimal computational overhead. This makes it ideal for lightweight wearable devices connected wirelessly to a multi-sensor network. Furthermore, the model maintains a strong predictive performance even after significant feature reduction, thereby further reducing computational costs. The framework is extensively validated using real-world infant activity data collected over several days. Class prediction accuracies are evaluated across different sensor combinations and compared with a pre-trained supervised neural network model. Results show that the proposed SSL framework achieves up to 80% accuracy in classifying multiple classes of infant activities (e.g., movement, sedentary behavior, falls) when using data from combination of all the three sensors in the system. Additional analysis explores the relationships between prediction accuracy and computational overhead with varying sensor count, feature reduction and clustering algorithms during the learning phase.**

Keywords*— *multi-sensor framework; semi-supervised learning; on-device learning; personalized learning; infant activity detection

I. Introduction

The objective of this paper is to develop a framework for semi-supervised learning of infant activity detection using multi-data streams from multiple sensors. Many sensor-based health and wellness monitoring applications such as smart hydration tracking [1] and human activity detection [2] use supervised classification algorithms using a large quantity of pre-labeled data. Due to privacy and other practical concerns, when such a large amount of pre-labelled data becomes unavailable, semi-supervised learning can provide an alternative solution. This approach usually relies on a small amount of data that can be categorized (i.e., labeled) into specific classes based on determinant feature values of the corresponding sensor data.

Semi supervised learning and activity classification can be performed on-wearable devices or out of devices depending on the application demands. Data privacy requirements in many applications call for on-device learning, which in turn requires learning algorithms to be run on resource-constrained device platforms. This calls for developing learning mechanisms with low computational overhead. The challenge of developing semi supervised algorithms with less computational overhead becomes more compounded when the learning and classification involves data from multiple sensor sources, which is the scenario handled in this paper. More specifically, we deal with an infant activity classification task that collects data from three on-body sensors and run semi-supervised learning algorithms for identifying activities.

Motor skill development in infants is known to have large subject-specificity for monitoring the wellness among infant subjects. Thus, generalized supervised learning seems less appropriate compared to subject-specific personalized learning. To that end, we deploy a semi-supervised learning model that trains itself with sensor data from an infant, whose activities are subsequently classified by the model. This allows the model to learn any unique nuances present in the infant's motor activities, thus rendering a personalized classification model. This incorporates a natural adaptability of the model in that it can continuously retrain to capture any changes in the activities of the infant over time.

The paper deals with a multi-sensor approach to infant wellness monitoring through activity detection for the following reasons. Unlike the usual adult activities, many primary activities of an infant including lying down, crawling, cruising etc. do not involve a straightforward upright torso configuration. This, coupled with the fact that infant activities fundamentally suffer from some lack of uniformity, adds many nuances in how the body moves for an infant. As a result, solutions [2] based on a single accelerometer and/or orientation sensor often prove to be insufficient for high resolution activity classification for the infants. Instead, a three-sensor approach is used in which the sensors are attached on the chest, thigh, and the upper arm of the subject infants. Data from all the sensors are consolidated in a single sensor, in which the proposed semi-supervised learning algorithms are executed for activity classification.

The proposed learning framework was experimentally analyzed and validated using data involving 12 hours of infant activities collected from two infants over a duration of several days. The infant subjects spontaneously performed their regular activities in an unconstrained environment. The resulting motion data were collected from the three accelerometer sensors mentioned above. It has been shown that the proposed semi-supervised learning-based classifier is able to classify infant activities (*i.e., moving, sedentary and fall*) with an accuracy up to 80%. To demonstrate the light-weight nature of the deployed algorithms, their computational complexities are also analyzed against the corresponding classification accuracies.

The paper has the following scope and contributions. First, an semi-supervised learning (SSL) model for a multi-sensor system has been developed for infant activity classification. Second, a detailed characterization was performed in order to understand the impacts of different sensors and their combinations on classification performance and computational complexity. Third, the SSL model was experimentally explored further to understand its performance with varying degrees of feature dimension reduction. Finally, all steps of the model execution were characterized for evaluating their computational overhead like time and memory usage, which is relevant for running the model and embedded hardware platforms.

II. System and Data Model

A. Sensing Modality and Dataset

As depicted in Fig 1, three wearable accelerometer sensors have been fitted on three parts of an infant's body, namely, the chest (S1), the upper arm (S2), and the thigh (S3). Each of these sensors are small (i.e., 3.75×1.5×0.5 inches) and lightweight (approximately 55 grams). The sensors were placed in custom-designed pockets to ensure snug fit and minimal discomfort. All three sensors record acceleration data with sampling rate of 50Hz. Data from the sensors were collected from two subjects and for 12 hours over a period of several days. Six different activities including crawling, cruising, standing, sitting, walking, and falling were recorded for both the subjects.

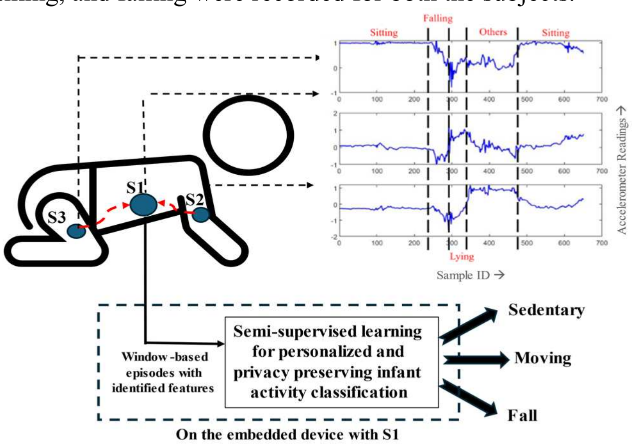

Fig 1: System architecture of the proposed multi-sensor framework-based infant activity detection system

The real-time acceleration data recorded by S2 and S3 are transmitted via a wireless network to S1. Subsequently, the data pre-processing and semi-supervised learning (SSL)-based classification process is carried out in the embedded sensor device S1. This is done using the combined data from all three sensors. The SSL model gets trained based on the real-time data of the infant subject, and subsequently detects three broad classes of infant activity states namely, *moving, sedentary, and fall*.

B. Pre-processing and feature engineering

The data read by an accelerometer sensor is in a time-series. A single accelerometer can read motion in three-orthogonal directions (*x, y, z*). Thus, the recorded data from a single accelerometer is in the form of three series. There are three accelerometers used for this proposed framework, so the classifier model will be working with 9 time series (*3 accelerometers * 3 time-series/accelerometer*). The time series are segmented into windows of 0.5 second worth of data, which means 25 samples (*50Hz * 0.5s*) from each time series. Therefore, each window consists of 225 samples of data (*25 samples * 9 time-series*). Each window constitutes a data point for the classifier model. Each datapoint is normalized by subtracting the mean value from the readings in each of the 9-time series. The next step involves feature extraction of each datapoint. Two features from each of the 9 time-series have been extracted. The first one is the coefficient of variation, which depicts the variability of acceleration across 25 samples spanning the 0.5s long window. The second feature is the number of mean crossing points within a 0.5 second window. This feature physically indicates the frequency of movements of the subject from its mean position. Thus, each datapoint is represented by 18 features (2 features * 9 time-series). These 18 features are fed into the proposed Semi-supervised learning-based classifier for simultaneously self-training and predicting activities of the three classes outlined next.

C. Class definition and distribution

In various types of infant movements including crawling, cruising, standing, sitting, walking, and falling, we aim to detect three broad movement classes namely, *moving, sedentary, and fall.* These classes are chosen based on the need for understanding infants' ambulatory motor development. The collected data that is used in this work has 3600, 6067 and 632 datapoints from the moving, sedentary and fall classes, respectively. Since the data was collected for infant subjects in their natural home setting, the data distribution across the classes was not uniform. A K-means clustering-based Synthetic Minority Oversampling Technique (SMOTE) [3] was used to balance the dataset after which we have 6067 datapoints from each of the three classes. Unless otherwise stated, this balanced dataset is used for all the reported work in the rest of the paper.

D. Processing Pipeline

Complete processing pipeline of the proposed semi supervised learning framework is depicted in Fig 2. The 50 Hz motion data recorded by the three different sensors are assembled in the memory on the embedded device of sensor 1 (S1). The 9 time series data are segmented into a 0.5 second window which constitutes a datapoint (episode). Each datapoint is normalized. Two features from each of the 9 time-series are extracted. The extracted features represent each episode which are unlabeled. Among the unlabeled episodes with identified features, some episodes are pre-labelled into 3 classes based on some extreme feature values beyond a threshold. These are considered pre-labelled because they can be easily identified as the parent of a class based on their extreme feature values matching to the corresponding class. These pre-labelled episodes are used in the SSL training phase for the cluster labelling step. All processing operations can be carried out on a lightweight embedded device with computational constraints. To that end, reducing the data dimension can lead to reduced computation and other resource loads. Popular dimensionality reduction techniques such as Principal Component Analysis (PCA) can be applied on the identified features of the unlabeled dataset in order to select the most discriminatory features. The unclassified (unlabeled) episodes with PCA-reduced feature-set are used by the proposed iterative semi-supervised learning model to self-learn and classify the movements episodes into the target three classes.

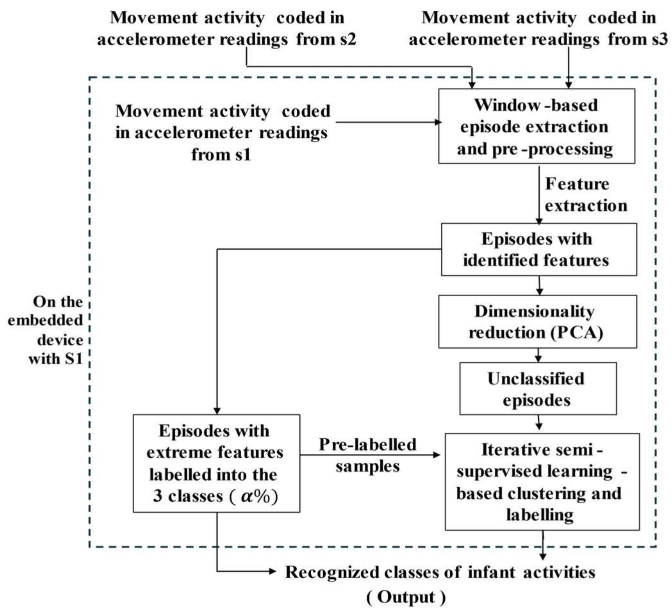

Fig 2: Processing pipeline of the proposed learning framework.

The pre-labelled episodes cater to the labelling process in the semi-supervised learning environment. As the infant keeps on performing ambulatory activities, new undetected episodes keep on accumulating in the unlabeled data pool. The iterative semi-supervised learning model runs iteratively after a certain number of incoming episodes. This processing pipeline enables continuous improvement in the real-time infant activity recognition accuracy over time, with half-a-second temporal granularity.

III. Semi-supervised learning-Based Multi-Sensor Infant Activity Detection framework

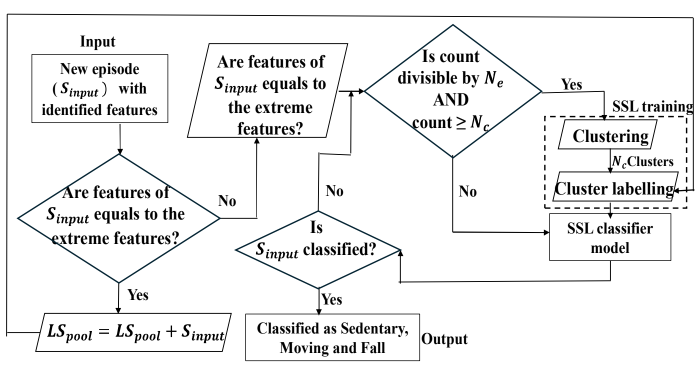

Fig 3: Iterative semi-supervised learning framework

The proposed iterative semi-supervised learning framework [4], [5] is depicted in Fig 3. As an incoming episode (pre-processed and normalized) (S_{input}) defined by the reduced feature set arrives, it is added to the data pool (S_{pool}). It is also checked whether any of the feature values of the new incoming episode is beyond a certain feature boundary, that particular episode is added to the pre-labelled data pool (LS_{pool}). After every N_e number of incoming episodes, the SSL training process takes place on the present S_{pool}. The training process involves two steps. The first being clustering the episodes in S_{pool} into N_C number of clusters. Varieties of K-means or Gaussian Mixture Model (GMM) clustering algorithms are used in this step. The N_c clusters are labelled using the pre-labelled data pool (LS_{pool}). A population-based labelling technique explained in algorithm 1 has been incorporated for the labeling of the clusters. In the population-based labelling, the cluster is labelled on the basis of the highest number of pre-labelled episodes present in that cluster. For example, if cluster $'a'$ has R_1 pre-labelled episodes from class X and R_2 pre-labelled episodes from class Y, and if $R_1 > R_2$, then cluster $'a'$ gets labelled as class X. If cluster $'a'$ does not have the highest number of pre-labelled episodes class of pre-labelled episodes or does not have a pre-labelled episode at all, it remains unlabeled. Thus, all the episodes in that cluster remain unclassified. All the episodes belonging to a cluster get labelled according to the label of the cluster.

Algorithm 1: Population-based cluster labelling algorithm

1. For $i = 1\ to\ N_c$
2. Calculate the # of pre-labelled samples in a-th clusterR_1, R_2,…,R_n, for n classes.
3. if max (R) exists:
4. label cluster i with label of max(R)
5. else:
6. cluster a remains unlabeled.
7. End if-else
8. End for

After the clusters are labelled using the population-based cluster labelling algorithm, the labelled clusters act as the trained semi-supervised classifier model. When a newly arrived and unlabeled episode comes in between training iterations, the episode is classified based on the most recently trained model, which are the labelled clusters. The new episode obtains the class label of the nearest cluster in the feature space. It can be seen how the model with 40 clusters expands and the classification accuracy improves with more incoming episodes. The quantitative performance of the classifier is formally presented in the next section.

IV. Experiments and Results

A dataset consisting of 12 hours infant activity data collected from 2 infant subjects over a period of several days have been used for validation of this proposed multi-sensor Semi-supervised learning (SSL) framework. The movement data of the infant subjects were recorded from three sensors fitted on three parts of the infants' body viz *chest, arm and thigh* as discussed in section III-A. Time-stamped video data was also recorded to get the ground truth information of the ambulatory activities performed by the concerned infant subject for the validation purpose of the learning model. The raw time-series data from the 3 accelerometers has been windowed and pre-processed to extract episodes (datapoints) of 0.5 seconds length as mentioned in section III-B. 6 features from each accelerometer sensors' data have been extracted to represent an episode for the proposed learning model which is mentioned in section III-C. The proposed learning framework discussed in section IV has been developed to classify among three classes of infant activities: *moving, sedentary and fall.* The degree of precision to classify the mentioned classes by proposed framework is evaluated by means of 3 parameters: *True positive (tp) for each class, false positive (fp) for each class and the overall accuracy.* In order to analyze the feasibility of the proposed framework for a wearable embedded device, *computational time (s) and CPU memory usage (Bytes)* have been considered as the representation of the computational overheads of the framework.

A. *Pre-trained supervised learning model (benchmark) vs. proposed semi-supervised learning framework*

A pre-trained supervised model (Artificial Neural Network-NN) evaluated with a 10-fold cross-validation using the data from the combinations of different sensors fit same dataset has been used as a benchmark for the proposed semi-supervised learning framework. The performance of the benchmark and the proposed learning algorithm are evaluated using data from the different combinations of 3 sensors. The number of input features to the learning model varies according to the data from the number of sensors used. 6 extracted features come from each of the sensors' data as mentioned section III-C. Thus, there are 6 input features for one sensor combination, 12 features for 2 sensors' combination and 18 features for 3 sensors combination (*sensor 1: on chest, sensor 2: on arm, sensor 3: on thigh*). The NN model is trained using the hyper-parameters mentioned in table 1.

TABLE 1: HYPER-PARAMETERS FOR PRE-TRAINED NN (SUPERVISED LEARNING) MODEL

# of input features	6/12/18 (based on the number of sensors)
# of hidden layers	9 hidden layers (36, 54, 108, 216, 432, 200, 100, 50, 25 neurons respectively)
# of output classes	3
Loss function	Categorical Crossentropy
Optimizer	Adam
Activation function	Relu (hidden layers), sigmoid (output layer)

The hyper-parameters of the proposed SSL framework include the number of input features. The input features in consistence with the number of sensors used. The SSL model has 40 clusters (N_c), 15% of pre-labelled datapoints (α).The cluster model gets trained after every 100 incoming datapoints (N_e), as read by the accelerometers. K-means clustering is used in the training process of the proposed SSL. The class-wise *tp* and *fp* of both NN and the SSL (post-convergence) are depicted in the heatmaps in Fig 4 (a-d). Classes 1 and 2 (moving and sedentary) are detected with higher *tp* and less *fp* values for all sensor combinations in NN as compared to SSL. Both of these classes are detected with approximately 80-88% *tp* and 5-12% *fp* by a pre-trained NN model, when any combination involving sensor 3 (on thigh) (S3, S2+S3, S1+S3, S1+S2+S3).

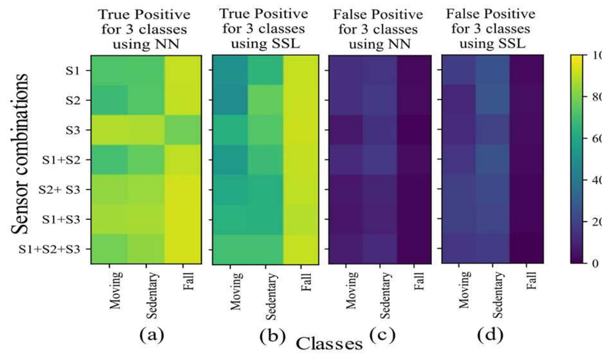

Fig 4: Class-wise true positive and false positive of NN and SSL.

With similar sensor combinations, SSL also performs better than the other sensor combinations for classes 1 and 2. But the *tp* (~65-75%) and *fp* (~12-20%) values are slightly worse compared to same with NN. Class 3 (Fall) is always detected with very high *tp* (around 90%) and low *fp* (around 1%) by both NN and SSL, with all the sensor combinations. Fig 5 compares the overall accuracy of the pre-trained NN and the proposed SSL frameworks. The figure shows that SSL has the best overall accuracy (78%) when all 3 sensors are used (S1+S2+S3) and the performance is very close to that of the pre-trained NN (85%). SSL also works better than other sensor combinations in terms of the overall accuracy (75%), when only sensor 3 is used. This overall accuracy value of SSL is also compared to what we get using a pre-trained NN (82%) with the same sensor (sensor 3).

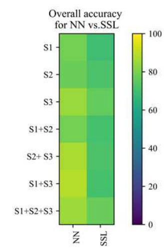

Fig 5: Overall accuracy of pre-trained NN vs proposed SSL.

B. *Impact of clustering algorithms in SSL with different sensor combinations*

Fig 6 (a-d) analyzes the post-convergence accuracy (*tp, fp for classes 1-2, and the overall accuracy*) for the proposed SSL algorithm with different cluster algorithms and sensor combinations. The clustering algorithms considered here are three variants each of K-means clustering and Gaussian Mixture model (GMM) [9]. K-means clustering methods include plain vanilla K-means (KM) [6], Mini-batch K-means (MBKM) [7], Bisecting K-means (BiKM) [8] while GMM includes the different covariant types of the clusters viz spherical (GMMs), full (GMMf) and diagonal (GMMd).

It can be observed in Fig 6(a), Class 1 has the best *tp* when 3 sensors are used all together (S1+S2+S3) along with GMMf clustering (around 88%). But the *fp* for the similar set-up is also

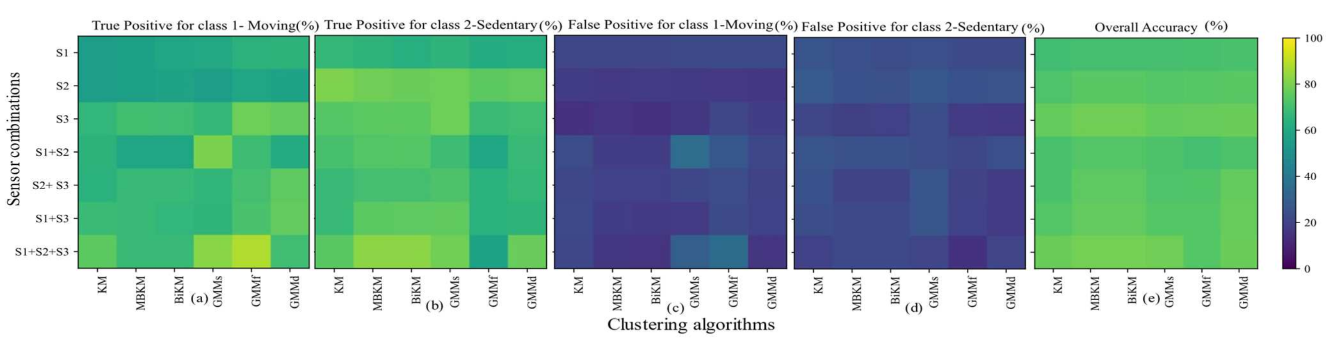

Fig 6: (a)-(b) True positive; (c)-(d) False positive for classes 1 and 2 and (e) Overall accuracy for SSL using different clustering algorithms and sensor combinations.

very high (around 30%) as shown in Fig 6(c). The *tp* for class 2 is low (around 70%) and the *fp* is (>30%) for the same class (shown in Fig 6(b), (d)). BiKM and MBKM have very high *tp* (>82%) for class 2 along with low *fp* (<15%) when 3 sensors are used (S1+S2+S3) but both of their *tp* for class 1 is comparatively low ie ~67%. GMMs has a very good *tp* and *fp* for both class 1(*tp* ~78%; *fp* ~25%) and class 2 (*tp* ~82%; *fp* ~20%), when all the 3 sensors are used. Class 3 has a very high *tp* (~92%) and low *fp* (~1%) for all the sensor combinations and the clustering algorithms. Thus, it can be observed that GMMs, MBKM and BiKM have the best overall accuracy (~80%) for predicting the 3 infant activity classes using motion data from 3 sensors (as shown in Fig 6 (e)). It can also be observed that S3 also has a good overall accuracy (~78%) with the same clustering algorithms. Thus, it is quite evident that data only from S3 has a better prediction accuracy than any other 2 sensors' combinations.

Fig 7 (a-b) depicts that using higher number of sensors means higher number of input features to SSL model, thus require higher computational overload (computational time and CPU memory usage) to train 18000 datapoints. Although GMMf and GMMs have very high overall accuracy, the CPU memory usage is very high for the GMM-based clustering algorithms (~2-2.5KB) for all the sensor combinations. Thus, BiKM and MBKM seem to be a better choice as a clustering algorithm for the proposed SSL algorithm on a wearable device as they have a very good prediction accuracy (~80%) consuming comparatively less computational overhead (*time*: ~19hrs; *CPU*: ~200 Bytes) using 3 sensors for simultaneously training and predicting 18000 datapoints.

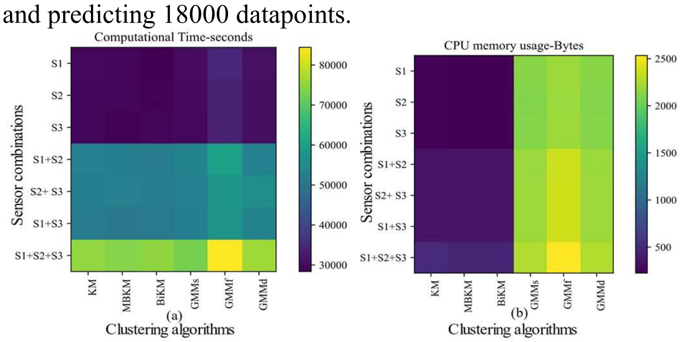

Fig 7: Computational overhead results for varying sensor combinations and clustering algorithms. (a) Computational Time(s); (b) CPU memory usage (B)

C. *Impact of feature reduction on proposed multi-sensor SSL performance.*

Feature reduction techniques like Principal component analysis (PCA) help in reducing input features to the learning model and thus reducing the computational load of the learning paradigm on the wearable device fitted on the infant's body. The challenge is to reduce the number of features to such an extent till the prediction accuracy of the learning model is preserved. Fig 8 (a) and (b) analyze the effect of the number of input features on the overall accuracy and the computational time of the proposed multi-sensor SSL framework using data from different number of sensors. The feature reduction analysis is done on the basis of the results involving this Bisecting K-means (BiKM) clustering algorithm. While using data from 3 sensors, the accuracy (~80%) is preserved till the number of features are reduced from 18 to 10 (45% feature reduction). Also, for 2 sensors' data, it can be observed that the accuracy (~75%) is maintained till the number of features are reduced from 12 to 9 (33% feature reduction). Thus, the higher the number of sensors, better accuracy is achieved which is preserved with higher percentage of feature reduction. Also, till the point where the accuracy is maintained for a particular number of sensors, it can be observed that with same number of features, data from higher number of sensors have better accuracy in less computational time as compared to that with the data from a smaller number of sensors. For example, when the number of features is 10 and 12, the proposed SSL with the 3-sensors' data has a better accuracy and less computational time than the proposed learning algorithm with 2-sensors' data (encircled with brown in fig 8 (a-b)).

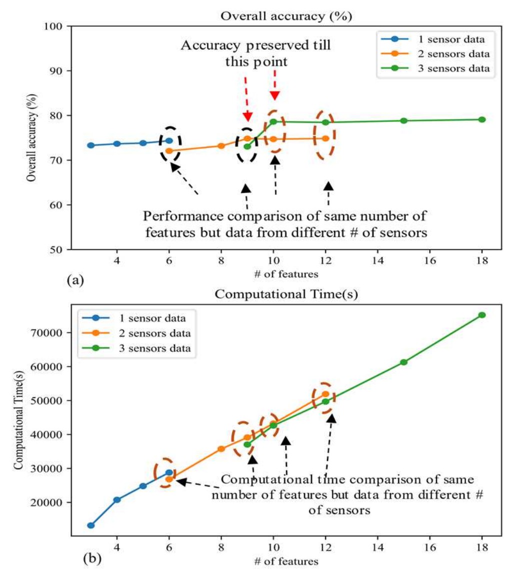

Fig 8: (a) Overall accuracy and (b) Computational time for different number of features used for the proposed SSL framework with different sensors.

This is due to the reason that 3 sensors' data has more information as compared to the 2-sensors' data, thus the cluster models converge faster for the former, thus consuming marginally less learning time. Since accuracy of the proposed SSL with 3 sensors' data decreases with the number of features below 10, the accuracy also falls below that of same algorithm with 2-sensors' data (encircles with black in Fig 8 (a)). Thus, the proposed SSL with 3 sensors' data and 45% reduced features will provide the best accuracy (~80%) with least possible computational overhead (~11 hours' *time*) while iteratively learning and predicting 18000 datapoints.

V. Related Work

There are very popular real-time personalized wellness monitoring applications like human activity detection which use very common inertial measurement unit (IMU) sensors like accelerometers and gyroscopes. Here the classification task is carried out by pre-trained supervised learning models on smartwatches and smartphones [10]. Different combinations of sensor modalities and their positions on the body have been experimented with in many existing works [2]. Works that use pre-trained supervised learning model to identify the infant activities [11-12] leverages multiple-sensor framework fitted on an infant body on boy parts like chest, arm and thigh. The primary drawback of such approaches is the usage of multiple learning models which are not-suitable for a lightweight on-device scenario with computational constraints. Using a single multi-class classification learning model would be more suitable for this scenario.

HAD applications usually deal with activities which are mostly subject-specific. This requires person-specific pre-labelled data, which can be scarce. Some existing human activity detection (HAD) systems uses a variety of pre-trained supervised learning-based classifiers [13-14] trained on a generalized dataset collected from a variety of subjects. This negates the person-specific angle to these applications. Semi-supervised learning (SSL) [15] is a possible solution for on-device learning in which classifiers can be trained on the device from a specific target user's data, thus preserving the personalization feature of the model as discussed in [16]. In this paper, a semi-supervised learning algorithm with sparsely labelled datapoints have been implemented. This same learning framework have been implemented on a binary [4] and a multi-class classification [5] scenario. This mechanism has proven to be effective for lightweight on-device applications like the above. An SSL method based on the combination of deep learning and transfer learning has been implemented in [17]. Usage of pre-labelled training data from different users in a deep learning model is required which affects the person-specific feature of this application. The work in [18] presents a HAD system using a bi-view SSL framework to detect semantic human activities like having dinner, shopping, etc. Windowed datapoint extraction technique and clustering mechanism have been implemented here as a classifier model, but externally on a different device. This method also involves a two-layered framework for the classification task. making it computationally expensive. Thus, it is not very suitable for an on-device self-training and classification solution.

The above works in the literature discusses many useful semi-supervised learning solutions but all with the usage of a single sensor. Applications like that of an infant activity detection are personalized. Also, due to a combination of non-upright and upright torso orientations, a framework of multiple sensors is required as well. Thus, a combined framework consisting of a combination of multiple sensors and a semi-supervised learning approach can be an effective solution approach for identifying multiple classes of infant activities. This paper sets out to accomplish that.

VI. Summary and Conclusions

This paper deals with a wireless multi-sensor approach to infant wellness monitoring by activity classification using a novel on-device semi-supervised learning. Data from multiple sensors are aggregated in a single sensor, in which the proposed semi-supervised learning (SSL) algorithms are executed. The proposed SSL paradigm can simultaneously train and predict multiple classes of infant activities (moving, sedentary, fall) without the presence of any pre-labelled data with the expense of least possible computational overhead. Thus, it is suited for a lightweight wearable device. The proposed learning framework was experimentally analyzed and validated using data involving several hours of infant activities for different sensor combinations and the results were compared to the results obtained by a pre-trained supervised neural network model. It has been shown that the proposed semi-supervised learning-based classifier is able to classify infant activities with an accuracy up to 80% when all the 3 sensors have been used. Further analysis has been done based on prediction accuracy and computational overheads of the proposed SSL framework using different clustering algorithms for different sensor combinations. Lastly, it has been observed that data used from all the 3 sensors in the proposed learning framework sustains its performance for reduction in feature dimensions of up to 45% which helps in further cutting down the computational overhead of the proposed learning paradigm.

VII. Acknowledgement

This work was funded by a grant (Award No.: 2427573) from the National Science Foundation.

References

[1] A. Roy, H. Dutta, H. Griffith, and S. Biswas, "An On-Device Learning System for Estimating Liquid Consumption from Consumer-Grade Water Bottles and Its Evaluation," *Sensors*, vol. 22, no. 7.
[2] N. Twomey *et al.*, "A Comprehensive Study of Activity Recognition Using Accelerometers," *Informatics*, vol. 5, no. 2, Art. no. 2, Jun. 2018.
[3] Z. Xu, D. Shen, T. Nie, Y. Kou, N. Yin, and X. Han, "A cluster-based oversampling algorithm combining SMOTE and k-means for imbalanced medical data," *Inf. Sci.*, vol. 572, 2021.
[4] A. Roy, H. Dutta, A. K. Bhuyan, and S. K. Biswas, "Semi-Supervised Learning Using Sparsely Labelled Sip Events for Online Hydration Tracking Systems," in *2023 International Conference on Machine Learning and Applications (ICMLA)*, 2023, pp. 1799–1804.
[5] A. Roy, H. Dutta, A. K. Bhuyan, and S. Biswas, "On-Device Semi-Supervised Activity Detection: A New Privacy-Aware Personalized Health Monitoring Approach," *Sensors*, vol. 24, no. 14, Art. no. 14, Jan. 2024.
[6] K. P. Sinaga and M.-S. Yang, "Unsupervised K-Means Clustering Algorithm," *IEEE Access*, vol. 8.
[7] M. M. Chavan, A. Patil, L. Dalvi, and A. Patil, "Mini Batch K-Means Clustering On Large Dataset".
[8] J. Di and X. Gou, "Bisecting K-means Algorithm Based on K-valued Self-determining and Clustering Center Optimization," *J. Comput.*, 2018.
[9] Y. Zhang *et al.*, "Gaussian Mixture Model Clustering with Incomplete Data," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 17, no. 1s, Mar. 2021.
[10] A. Rasekh, C.-A. Chen, and Y. Lu, "Human Activity Recognition using Smartphone," Jan. 30, 2014, *arXiv*: arXiv:1401.8212.
[11] H. Dutta, A. K. Bhuyan, M.-H. Lee, and S. Biswas, "Towards a Multi-Sensor Paradigm for High Resolution Infant Activity Classification," in *SoutheastCon 2024*, 2024.
[12] R. K. Thelagathoti *et al.*, *AIARS: Towards Automated Infant Activity Recognition System using Machine Learning and Multi-sensor Fusion Data*. 2024.
[13] Y. Wang, S. Cang, and H. Yu, "A survey on wearable sensor modality centred human activity recognition in health care," *Expert Syst. Appl.*, vol. 137, 2019.
[14] H. F. Nweke, Y. W. Teh, M. A. Al-garadi, and U. R. Alo, "Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges," *Expert Syst. Appl.*, vol. 105, 2018.
[15] M. F. A. Hady and F. Schwenker, "Semi-supervised Learning," in *Handbook on Neural Information Processing*, M. Bianchini, M. Maggini, and L. C. Jain, Eds., in Intelligent Systems Reference Library., Berlin, Heidelberg: Springer, 2013.
[16] G. Singh, M. Chowdhary, A. Kumar, and R. Bahl, "A Personalized Classifier for Human Motion Activities With Semi-Supervised Learning," *IEEE Trans. Consum. Electron.*, vol. 66, no. 4, 2020.
[17] S. Oh, A. Ashiquzzaman, D. Lee, Y. Kim, and J. Kim, "Study on Human Activity Recognition Using Semi-Supervised Active Transfer Learning," *Sensors*, vol. 21, no. 8, Art. no. 8, Jan. 2021.
[18] M. Lv, L. Chen, T. Chen, and G. Chen, "Bi-View Semi-Supervised Learning Based Semantic Human Activity Recognition Using Accelerometers," *IEEE Trans. Mob. Comput.*, vol. 17, no. 9, Sep. 2018.