SPHERE-DNA: Privacy-Preserving Federated Learning for eHealth

Jari Nurmi
EE Unit
Tampere University
Tampere, Finland
jari.nurmi@tuni.fi

Yinda Xu
EE Unit
Tampere University
Tampere, Finland
yinda.xu@tuni.fi

Jani Boutellier
School of Technology and Innovations
University of Vaasa
Vaasa, Finland
jani.boutellier@uwasa.fi

Bo Tan
EE Unit
Tampere University
Tampere, Finland
bo.tan@tuni.fi

***Abstract*—The rapid growth of chronic diseases and medical conditions (e.g. obesity, depression, diabetes, respiratory and musculoskeletal diseases) in many OECD countries has become one of the most significant wellbeing problems, which also poses pressure to the sustainability of healthcare and economies. Thus, it is important to promote early diagnosis, intervention, and healthier lifestyles. One partial solution to the problem is extending long-term health monitoring from hospitals to natural living environments. It has been shown in laboratory settings and practical trials that sensor data, such as camera images, radio samples, acoustics signals, infrared etc., can be used for accurately modelling activity patterns that are related to different medical conditions. However, due to the rising concern related to private data leaks and, consequently, stricter personal data regulations, the growth of pervasive residential sensing for healthcare applications has been slow. To mitigate public concern and meet the regulatory requirements, our national multi-partner SPHERE-DNA project aims to combine pervasive sensing technology with secured and privacy-preserving distributed privacy frameworks for healthcare applications. The project leverages local differential privacy federated learning (LDP-FL) to achieve resilience against active and passive attacks, as well as edge computing to avoid transmitting sensitive data over networks. Combinations of sensor data modalities and security architectures are explored by a machine learning architecture for finding the most viable technology combinations, relying on metrics that allow balancing between computational cost and accuracy for a desired level of privacy. We also consider realistic edge computing platforms and develop hardware acceleration and approximate computing techniques to facilitate the adoption of LDP-FL and privacy preserving signal processing to lightweight edge processors. A proof-of-concept (PoC) multimodal sensing system will be developed and a novel multimodal dataset will be collected during the project to verify the concept.**

***Index Terms*—multi-partner project, machine learning, differential privacy, LDP-FL**

I. Introduction

With the overall general tendency of population ageing and simultaneous shortage on personnel in the healthcare sector, use of various types of sensors in residential environments for modelling resident activity and consequently inferring potential health or wellbeing issues of residents has been studied widely [1]. Successful modelling and detection of resident activities can reveal information related to active and rest periods, walking patterns, possible hand/arm tremor and body gestures that can be signals of chronical medical conditions such as depression, Parkinson's disease, diabetes, dementia and so forth. Furthermore, instant and accurate fall detection (possibly incurred by a stroke) can also improve emergency response and patient recovery time. Previous research like SPHERE IRC [2] in the University of Bristol has proved that combining data originating from a variety of home sensors can accurately characterize resident activities. Leveraging advanced and widely deployed ICT infrastructure, researchers have been able to aggregate large volumes of multiple sensor data from a number of households, aggregating it to a central cloud for activity modelling and/or deep learning. Despite the promise of improving wellbeing, use of residential sensor data for healthcare purposes also raises serious concerns related to privacy and data security due to the use of imaging, voice recording, potential security compromises of data centers and inverse reasoning. Also, misdiagnosis caused by false detections or inaccurate models is a concern when relying on automated sensing for healthcare.

Our SPHERE-DNA project proposes advancing the state-of-the-art in privacy-preserving human action recognition by a holistic, machine learning (ML) based approach. In the project, the best combination of sensor types and security architectures is determined by a ML algorithm from a set of carefully selected sensing and security technologies, combined with edge computing expertise and field tests. The three-partner national project (two groups from Tampere University and one from University of Vaasa) started in January 2022 and runs for three years.

The rest of the paper is organized as follows. In Section II we will present the state of the art in this area. Section III will detail the proposed work, whereas Section IV presents some preliminary results and Section V concludes the paper.

II. State-of-the-Art

In this section, we present the state-of-the-art in the key technologies for privacy-preserving health monitoring in residential environment.

The Academy of Finland is supporting this work via the SPHERE-DNA project (No. 345681) in the ICT 2023 programme.

978-3-9819263-7-8/DATE23/© 2023 EDAA

A. Data privacy for distributed healthcare systems

To address security, privacy and misdiagnosis concerns, researchers have proposed various risk mitigation strategies. The main trend is to decentralize data, algorithms and models to devices ran by multiple parties, also applying privacy preserving processing on raw and processed data [3]. Amongst the ongoing works, federated learning (FL), differential privacy (DP), on-device learning and privacy preserving sensing (especially for imaging) are promising techniques for achieving privacy preservation and accurate activity modelling. FL is a secure machine learning architecture, where sensitive raw sensor data is maintained in edge nodes (individual households or end devices) and used for local model weight updates. The updated distributed gradients are then sent to a centralized cloud for aggregation, thus avoiding transmission of sensitive data. DP, on the other hand, helps in protecting sensitive information by perturbation and noise injection before data is provided for external access.

B. Privacy preservation in imaging

Privacy-conscious imaging can be accomplished by a variety of approaches, which may or may not require specialized imaging devices. A straightforward approach is relying on very low-resolution image sensors [4], or alternatively, use of novel imaging sensor types or filters. Lensless cameras (or coded aperture) refer to types of imaging devices that yield visually unrecognizable image data, which still can be used for analytics purposes such as action recognition [5] by leveraging deep learning. An alternative to this approach is to use optics add-ons that obfuscate details, while preserving sufficient information for image analysis [6]. Finally, privacy preservation in imaging can also be achieved by algorithmically transforming the camera-captured image content into a representation that is not human intelligible. Recently, image transformation into line clouds [7] was proposed, making the depicted scenes unintelligible, but retaining sufficient information content for camera pose estimation.

C. Acoustic action recognition and localization

Given a network of acoustic sensors, the structure from sound problem aims at simultaneously localizing the sensors in the network as well as the acoustic events [8]. A typical structure from sound scenario consists of a network of synchronized receivers and unsynchronized transmitters. In this domain, research has focused on developing optimized algebraic solvers for so-called minimal configurations, i.e., configurations with the smallest number of nodes to ensure a finite number of solutions. This technique has been very successful in the computer vision domain, where it allowed to solve problems such as structure from motion or camera calibration within microseconds [9]. More recently, this approach has also been exported to the domain of sensor network calibration. While this approach allowed to fully solve calibrated networks [10], for unsynchronized cases only algebraic solvers for subminimal configurations have been available [11]. Recently, our researchers have been able to solve some minimal configurations for the 2D case [12].

D. On-device learning

Often machine learning algorithms are run on huge data centers. However, instead of centralized computations, it is possible to leverage hive intelligence by using decentralized computing near or on the sensor devices. This approach also supports the target of increased cybersecurity and decrease of (often unnecessary) communications towards a central cloud. Executing machine learning models on embedded devices is commonly known as Embedded Machine Learning [13] – TinyML refers to implementation of machine learning on very resource-restricted microcontrollers [13]. In particular, there are attempts to use Tensor Flow Lite in conjunction with RISC-V open-source processors, but tool integration is still not complete [14]. One promising approach for distributed ML inference is presented in [15], dividing the computing between the Edge and embedded devices. Researchers have also experimented with implementing storage-saving Posit arithmetic on RISC-V [16].

E. Cross-modal sensor fusion

Sensor fusion enables combining sensor readings originating from different sensor modalities (e.g. acoustic, radar, infrared) with the aim of increasing modeling and detection accuracy due to complementary information sources. Recently, sensor fusion has been increasingly performed by leveraging machine learning (e.g. [17]). For privacy preservation purposes, visual sensors (infrared or RGB camera) deployed for resident monitoring are frequently of very low resolution, however often trained with high-resolution data [18]. Notably, fusing of heterogeneous low- and high-resolution sensor data has also shown promise in pose and body gesture estimation [19], semantic and behavioral level modelling [20].

F. Shortcomings in the state-of-the-art

Although individual approaches such as FL, DP, privacy-preserving imaging, on-device learning and cross-modal data fusion provide a certain degree of security, privacy and accuracy for healthcare applications, the status quo still does not meet increasingly strict regulations (e.g. GDPR) and public expectations. For example, the study presented in [21] has shown that gradient sharing in FL is still a potential security issue even without direct access to raw data. DP, on the other hand, provides security safeguards due to its composability, robustness against post-processing, and graceful degradation in the presence of correlated data. However, DP requires a trusted party to host the data, which is inherently conflictive with distributed storing of residential data and against the philosophy of FL. The problems of privacy preserving imaging and image-based positioning include novel attacks that effectively compromise [22] privacy preservation schemes, which previously have been assumed to be secure [7]; high-definition visual data is on the other hand capable of very accurate activity detection, as far as gait characterization and

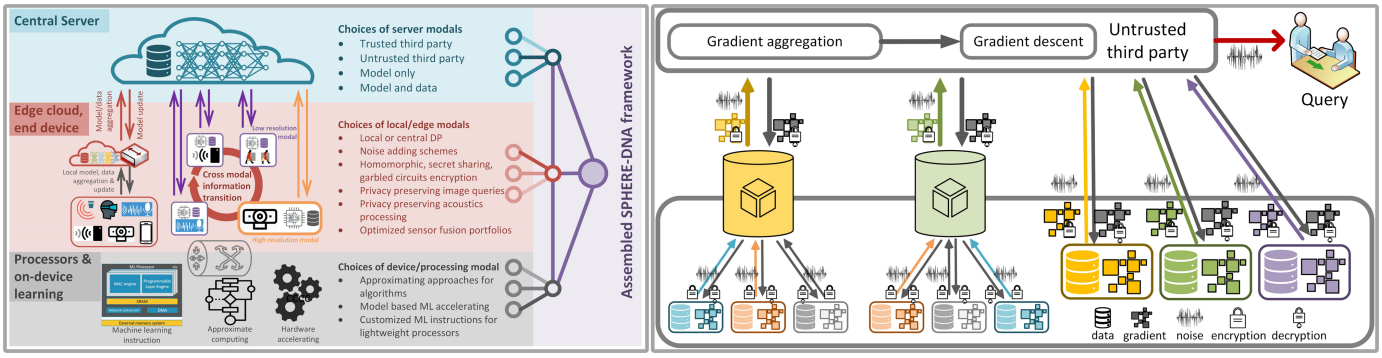

Fig. 1: Left: SPHERE-DNA project's structural approach for secure, privacy-preserving accurate activity modelling distributed with multi-modal data; Right: The schematic of the LDP-FL for the project's distributed multiple modality sensing data

vital sign (respiration) detection, however it can be also be used for identifying individuals from facial images even from extremely low resolution when machine learning is used [17]. Researchers have also attempted to achieve accurate human activity detection from fusing low-resolution sensing data from multiple modalities to avoid the use sensitive high-resolution imaging. Unfortunately, the accuracy achieved from fused low-resolution sensors is still lacking from the desired level. Related to the deployment of on-device learning, most of the current distributed sensors are only designed for collecting and transmitting data, with limited processing capacity and battery life, making them unfeasible for anything but very light signal processing, not to mention on-device learning. As a summary, the need for health-related activity monitoring in residential environments is unquestionable, however it is clear that the use of individual approaches for privacy-preserving sensing are not satisfactory but call for a holistic solution.

III. The Proposed Approach

Our multi-partner project proposes addressing secure and privacy-preserving human localization and action recognition by a holistic, machine learning based approach as shown in Fig. 1 Left. Our target is to advance the state-of-the-art in privacy-preserving activity modelling using i) individual sensor modalities (acoustic array, visual, infrared or radio frequency), but more importantly ii) we propose holistic, machine learning based sensor fusion and adoption of security mechanisms (FL, LDP). Our hypothesis is that by letting a machine learning based optimization process determine the best combination of sensor data modalities, as well as perform the exploration of security mechanism alternatives, it is possible to reach novel solutions and outperform the state-of-the-art in the achieved level of accuracy, while maintaining the desired level of security. Novel data security solutions are studied in terms of iii) flexible FL implementations, DP localization options, privacy preserving processing, and distributed computing architectures. For the purpose of implementing a machine learning based optimization process, our intent is also to study iv) performance metrics for security, privacy-preservation and activity recognition accuracy. Finally, we will v) develop a proof-of-concept system. The objectives of the project are as follows:

1) Convergence strategies for differential privacy and federated learning
2) Privacy preserving sensing mechanisms
3) Machine learning friendly power-efficient on-device processing capabilities
4) Integration and evaluation

We will now present in more details the content of the project to reach each of the objectives.

A. Convergence strategies for differential privacy and federated learning

1) Sub-objective 1.1: Computational and communications efficient federated learning: FL, proposed by Google [23], tries to protect privacy of the training data by sharing of gradient information only. This mechanism provides robustness against cyber attacks or server data breaches by keeping the sensitive data in the edge or user devices. However, the sensitive data can be still recovered from the shared gradients [24]. Thus, encryption methods are further applied to the shared gradients. These encryption strategies like secret sharing [25], homomorphic encryption [26] and obfuscated circuits [27], improve the user data security, while introducing high computational cost for encryption and increased transmission cost that harms edge computing solutions. Particularly, these approaches may require a trusted third party for the secret sharing, which is a potential security loophole. Also, most current FL models are designed for independent and identically distributed (i.i.d.) data, which is not always appropriate for multimodal sensor data that is hosted in different households or applied to the diagnosis of different types of health problems. In this context, our approach advocates to study BatchCrypt [28] type of federated learning strategies that are agnostic to network quality and processing capability, and suitable for non-i.i.d. multimodal data.

2) Sub-objective 1.2: Local differential privacy for time series healthcare data: Differential privacy (DP) [29] has been considered as an option for protecting the training data by adding an appropriate amount of noise to the data (perturbation). DP also works well with popular deep learning [30]

concepts such as DP-SGD. Compared to privacy preservation strategies like encryption (with high computational overhead) and anonymization (loss of original data), DP is almost an ideal privacy protection solution if there is a single trusted party who has access to the entire dataset. Thus, DP is often used for E-Health records of medical symptoms of a specific disease, or other clinical records in hospitals. With increasing numbers of smart home gadgets, a challenge is to protect the privacy of residential data, which can also be used to reconstruct personally identifiable information (PII). To establish robust privacy protection for household or individual sensor data, our aim is to design a flavor of local differential privacy that i) removes the need for a trusted third party, ii) relies on lightweight perturbation mechanisms, and allows iii) trading-off between data utility and privacy. Furthermore, we aim to study local DP (LDP) solutions for time series type of data by building the LDP solution embedding in the back-propagation through time (BPTT) process.

3) Sub-objective 1.3: Local differential privacy federated learning (LDP-FL): FL and LDP are countermeasures against active and passive attacks, respectively. This proposal envisions the combination of FL and LDP (Fig. 1 Right) for double resistance against data theft and inference attacks, as well as for flexibility to choose a suitable security architecture according to the local computing capability, network quality, and security needs. LDP-FL is expected to have features of active encryption, immunity against differencing and correlation attacks, low computing and transmission costs, and no need for a trusted third party. Different from DP, which aggregates raw or perturbated data, LDP-FL terminals share locally updated gradients between the edge and the central server. To avoid data leaks, encryption (for high-performance devices and connections), perturbation (for low-performance alternatives) or combinations of both methods (for users with high security requirements) will be implemented. However, there are open questions, such as how well the distributed gradient perturbation methods can perform in terms of balancing security, accuracy and complexity.

B. Privacy preserving sensing mechanisms

1) Sub-objective 2.1: Privacy preserving RGB imaging: As the previous examples show, hardware-based and algorithmic approaches that either capture visually unrecognizable content or algorithmically obfuscate recognizable content are subject to image reconstruction techniques that can compromise privacy if malicious users acquire access to sensor readings or obfuscated data. To this extent, the approach advocated for privacy preserving image analytics in this project is fundamentally different: we propose the use of edge computing such that the image sensor and the edge processor are decoupled from external network access. Consequently, a share of early visual processing needs to be performed on the edge device, similar to the neurosurgeon [31] approach. In this setting, the image sensor and its accompanied analytics processor only expose non-revertible statistical (or visual feature) information to the network interface, and consequently, accuracy-degrading design choices such as resolution reduction become unnecessary. This approach is studied as an alternative to federated learning.

2) Sub-objective 2.2: Acoustic array-based action recognition and subject localization: In the acoustic context, we will develop new algebraic techniques to solve sensor network calibration in 2D and 3D configurations. State-of-the-art algebraic approaches have followed a two-phase approach [9]: in the offline phase, heavy symbolic computations (e.g. Gröbner bases) are performed, and the code for a fast optimized solver is generated offline. In the online phase, the so-obtained solver can be used to quicky and accurately solve the system of polynomial equations. In this study we will integrate homotopy continuation, which is an algorithm from numerical algebraic geometry to solve problems in computer vision, into audio pipelines. The results of this study will 1) fully address the sensor network calibration problem from a mathematical perspective, and 2) use novel algebraic solvers to develop a fully functional real-time structure-from-sound pipeline.

3) Sub-objective 2.3: Cross-modal sensor fusion for activity recognition: In this objective, our intent is to use high-resolution video data to train a machine learning based activity recognition system that acquires data from various sources, such as infrared [32] and radar [33]. A promising approach are teacher-student networks [34] that are capable of transferring information from the video stream to low resolution modalities via keypoint confidence maps [35] as media, also connecting fragmentary characteristics from individual low resolution modalities. Particularly, under the considered LDP-FL configuration, gradient sharing and aggregation mechanisms need to be designed for distributed single-modal sensor nodes, model generalization and cross-modal information fusion on the server side.

C. Machine learning friendly power-efficient on-device processing capabilities

1) Sub-objective 3.1: Custom machine learning instructions for light-weight processors: In order to support on-device processing for distributed machine learning, the resource-constrained processor integrated in the edge device needs to be augmented by capabilities that increase its performance without leading to excessive power consumption. The target is to identify, implement and assess custom instructions for efficient processing of machine learning algorithms on the device, also building on the work in [36].

2) Sub-objective 3.2: Model-based approach for machine learning acceleration: Another approach for reaching the objective is to resort to an on-device machine learning accelerator to which the processing can be off-loaded from the lightweight processor. By using a model-based approach, the algorithms can be analyzed and synthesized onto the device. The scenario is to utilize reconfigurable logic capabilities on the device to build a tailored block to execute identified machine learning algorithm kernels.

3) Sub-objective 3.3: Approximate computing to lower on-device computation requirements: Complementary to the previous two approaches, the third sub-objective contributes

to efficiency by introducing approximate computing to custom instructions and/or reconfigurable hardware accelerator. Also, possibilities to save energy in other parts of sensor data processing will be investigated to unleash performance and power for intelligent privacy-preserving computing.

D. Integration and evaluation

1) Sub-objective 4.1: Unifying framework: This objective aims to build a framework for integrating the individual approaches (Objectives 1, 2 and 3) for complementary guarantees on privacy protection, flexibility in adapting to different computing and network configurations, and activity modeling accuracy. Our approach to the unified framework is a) constructing security-conscious distributed edge-cloud baseline architectures based on LDP-FL (Objective 1, Objective 2), and b) formulating multi-modal sensor fusion as a machine learning problem (Objective 2), which optimizes for activity recognition accuracy considering edge computing platform restrictions (Objective 3) as a cost function.

2) Sub-objective 4.2: Proof-of-the-concept system and dataset: Machine learning based activity recognition relies heavily on datasets. As Sub-objective 4.1 advocates machine learning based sensor fusion, there is a need to create a new dataset that covers all modalities considered in this proposal.

3) Sub-objective 4.3: Evaluation and metrics: For realizing Sub-objective 4.1, we will adopt and develop quantitative metrics for a) activity recognition performance, and b) computational cost. For both a) and b), standard metrics such as MFLOP/s, and classification accuracy exist, but they need to be adapted and formulated into a suitable cost function. Our intent is to treat the degree of adopted privacy safeguards (e.g., encryption, perturbation, privacy budget and sensitivity) as a user-given parameter that affects the adopted system architecture (see Sub-objective 4.1) and trades off the accuracy against computational complexity.

IV. Preliminary Results

A. Video-guided NLoS object tracking

We have developed a dual-modal (radio and optic) for indoor tracking and activity-capturing system without using imaging information sources. The system consists of a 3Tx-4Rx, 60∼64 GHz frequency-modulate continuous-wave (FMCW) mmWave radar (TI® IWR6843) and an Intel® RealSense L515 depth camera. The location inferred from RGB-D pixels is used as ground truth to train the radio sensing data with the lower concern of privacy leakage. In this work, the You only look once (YOLO) V4 is used for detecting and keeping track of the object.

Fig. 2 presents the Non-Line of Sight (NLoS) experimental scene. The 2D Multiple Signal Classification (MUSIC) algorithm is applied on the four receicving channels to estimate the range and angle of the strong reflections. The example rang-angle map (snapshot) is shown in Fig. 2. Due to complicated signal reflecting, the strongest spot on the range-angle map does not indicate the location of the person. Thus, a localization convolutional neural network (L-CNN) is used to train the

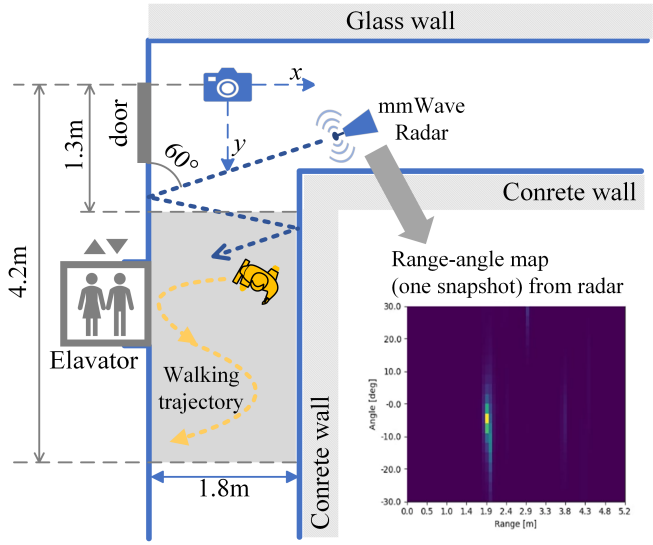

Fig. 2: The NLoS experimental layout diagram. The radar emits signals at a 60-degree angle to the wall and then collects bounced radio signals. The subject is walking in undefined trajectories within the shadowed area. The elevator and the door introduce complicating factors into our experiments.

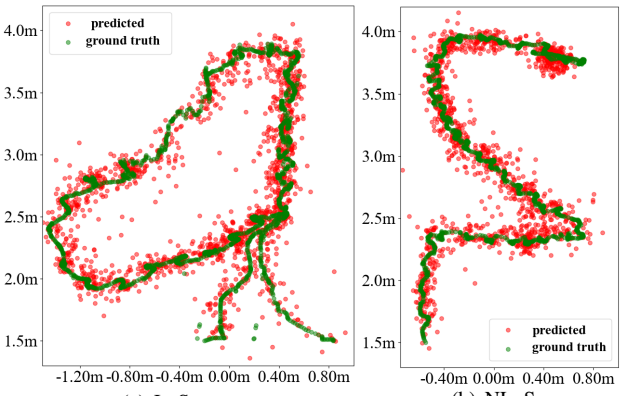

Fig. 3: The tracking results of LoS (a) scenario and NLoS (b) scenario (shown in Fig. 2). Green dots refer to ground truth points from RGB-D pixels, and red dots refer to the predicted results from the mmWave radio data.

snapshot by using the RGB-D based locations as ground truth. The L-CNN consists of seven 3D convolution layers. Each convolutional layer is followed by batch normalization and ReLU activation layers. Two fully connected layers are placed after the convolutional layers. The mean squared error (MSE) loss function during the training. The localization results based on the mmWave radio measurements and trained L-CNN are shown in Fig. 3 for both LoS and NLoS scenarios. The mean absolute error (MAE) of the LoS scenarios is (x, y) is $(0.0561\,m, 0.0701\,m)$. While the MAE of the NLoS scenario archives $(0.0818\,m, 0.0910\,m)$, which is an optimistic result with the presence of complicated layouts and misplacement of the radar and depth camera coordinations.

B. Privacy-preserving distributed object detection

We have developed a framework [37] that enables privacy-preserving inference for various machine learning based sensing applications such as object detection. Differing from previous works such as [31], our framework is based on a formally defined dataflow model of computing, and scales across a number of client (sensor) and server devices offering a selected degree of fault tolerance. The experiments carried out show promising results on the performance of the framework on embedded client devices and communication over wired and wireless communication channels. In future developments, the framework is to be extended from inference towards adopting the privacy-preserving training policies developed in this project.

V. Conclusions and future work

The overall hypothesis of SPHERE-DNA is that by jointly considering data privacy mechanisms (LDP, FL), multi-modal privacy preserving sensing, and realistic edge computing restrictions, it is possible to advance the state-of-the-art in privacy conscious activity recognition. More precisely, by unifying federated learning and local differential privacy, as well as edge computing in distributed sensor nodes, it is possible to provide complementing protection mechanisms against malicious data theft and inference attacks. Furthermore, formulating the sensor fusion problem as a machine learning problem enables the automatic discovery of optimal data source combinations and data features for accurate activity recognition. The preliminary results from the developed framework show the potential to act as the computing environment from our algorithms across to embedded sensor nodes and servers that assemble measurements from various inputs. The future work includes a designated neural network structure for multi-modal data fusion and decentralized training and noising strategies for multi-modal data from distributed sites.

References

[1] M. Vancea and J. Solé-Casals, "Population Aging in the European Information Societies: Towards a Comprehensive Research Agenda in eHealth Innovations for Elderly," Aging Dis. 2015, Dec 14;7(4):526-39.

[2] P Woznowski, et al. "SPHERE: A sensor platform for healthcare in a residential environment", in Designing, Developing, and Facilitating Smart Cities: Urban Design to IoT Solutions, Springer, 2017.

[3] Kaissis, G.A., Makowski, M.R., Rückert, D. et al. "Secure, privacy-preserving and federated machine learning in medical imaging" Nature, Machine Intelligence, 2, 305–311, 2020.

[4] Wang, Z., Miao, Z., Jonathan Wu, Q.M. et al. "Low-resolution face recognition: a review." Vis Comput, Springer, 2014.

[5] Z. W. Wang, et al., "Privacy-preserving action recognition using coded aperture videos." IEEE/CVF CVPR, 2019.

[6] F. Pittaluga and S.J. Koppal, "Pre-capture privacy for small vision sensors," IEEE TPAMI, 2016.

[7] P. Speciale, J.L. Schonberger, S.B. Kang, S.N. Sinha, and M. Pollefeys, "Privacy preserving image-based localization," IEEE/CVF CVPR, 2019.

[8] S. Thrun, "Affine structure from sound," in NIPS, 2005.

[9] V. Larsson, K. Åstrom, and M. Oskarsson, "Efficient solvers for minimal problems by syzygy-based reduction," in Proc. IEEE/CVF, CVPR, 2017.

[10] Y. Kuang, et al., "A complete characterization and solution to the microphone position self-calibration problem," IEEE ICASSP, 2013.

[11] S. Burgess, Y. Kuang, J. Wendeberg, K. Åström, and C. Schindelhauer, "Minimal solvers for unsynchronized TDOA sensor network calibration," in ALGOSENSORS, Springer, 2013.

[12] L. Ferranti, K. Åström, M. Oskarsson, J. Boutellier, and J. Kannala, "Sensor Networks TDOA Self-Calibration: 2D Complexity Analysis and Solutions," preprint arXiv:2005.10298, 2020.

[13] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures." IEEE Access, 2019.

[14] M. S. Louis, et al., Towards Deep Learning using TensorFlow Lite on RISC-V, Third Workshop on Computer Architecture Research with RISC-V, in CARRV, 2019.

[15] J. Boutellier, B. Tan, and J. Nurmi, "Fault-Tolerant Collaborative Inference through the Edge-PRUNE Framework," in Proc. International Conference on Machine Learning, ICML 2022.

[16] M. Cococcioni, F. Rossi, E. Ruffaldi, S. Saponara, Vectorizing posit operations on RISC-V for faster deep neural networks: experiments and comparison with ARM SVE, Neural Computing and Applications, Springer, 2021.

[17] M. Singh, S. Nagpal, R. Singh and M. Vatsa, "Dual Directed Capsule Network for Very Low Resolution Image Recognition," IEEE/CVF ICCV, 2019.

[18] M. Moencks, V. De Silva, J. Roche, A. Kondoz, Adaptive Feature Processing for Robust Human Activity Recognition on a Novel Multi-Modal Dataset. ArXiv, abs/1901.02858, 2019.

[19] M. Zhao et al., "Through-Wall Human Pose Estimation Using Radio Signals," IEEE/CVF CVPR, 2018.

[20] E. Kazakos, A. Nagrani, A. Zisserman, D. Damen, EPIC-Fusion: Audio-Visual Temporal Binding for Egocentric Action Recognition, IEEE/CVF ICCV, 2019

[21] L.T. Phong, Y. Aono, T. Hayashi, L. Wang, S. Moriai, "Privacy-Preserving Deep Learning via Additively Homomorphic Encryption," IEEE TIFS, 2018.

[22] K. Chelani, F. Kahl, and T. Sattler, "How Privacy-Preserving are Line Clouds? Recovering Scene Details from 3D Lines." arXiv:2103.05086, 2021.

[23] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B.A. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," Proc. 20th International Conference on Artificial Intelligence and Statistics, in PMLR 2017.

[24] L.T. Phong, Y. Aono, T. Hayashi, L. Wang, S. Moriai, "Privacy-Preserving Deep Learning via Additively Homomorphic Encryption," IEEE TIFS, 2018.

[25] K. Bonawitz, et al., P"ractical Secure Aggregation for Privacy-Preserving Machine Learning," in Proc. SIGSAC Conference on CCS, 2017.

[26] Q. Yang, Y. Liu, T. Chen, and Y. Tong. "Federated Machine Learning: Concept and Applications," ACM Trans. Intell. Syst. Technol. 2019.

[27] B. Rouhani, S. Riazi, F. Koushanfar. "Deepsecure: scalable provably-secure deep learning," ACM DAC '18. 2018

[28] C. Zhang, et al., "BatchCrypt: Efficient Homomorphic Encryption for Cross-Silo Federated Learning," USENIX ATC, 2020.

[29] C. Dwork, Differential Privacy, in International Colloquium on Automata, Languages, and Programming (ICALP), pp.1–12, LNCS, vol 4052. Springer, 2006.

[30] M. Abadi, et al., "Deep Learning with Differential Privacy," In Proc. ACM SIGSAC Conference on CCS, 2016.

[31] Y. Kang, J. Hauswald, C. Gao, A. Rovinski, T. Mudge, J. Mars, and L. Tang, "Neurosurgeon: Collaborative intelligence between the cloud and mobile edge," ACM SIGARCH Computer Architecture News, 45(1), 2017.

[32] Y. Karayaneva, S. Sharifzadeh, Y. Jing, K. Chetty and B. Tan, "Sparse Feature Extraction for Activity Detection Using Low-Resolution IR Streams," IEEE ICMLA, 2019.

[33] Y. Karayaneva, S. Sharifzadeh, W. Li, Y. Jing, and B. Tan, " Unsupervised Doppler Radar-Based Activity Recognition for e-healthcare," IEEE Access, 2021.

[34] Y. Aytar, C. Vondrick, and A. Torralba, "Soundnet: Learning Sound Representations from Unlabeled Video," in NIPS, 2016.

[35] G. Gkioxari, B. Hariharan, R. Girshick, and J. Malik, "Using k-poselets for Detecting People and Localizing Their Keypoints," In Proc. IEEE/CVF CVPR, 2014.

[36] S. Payvar, M. Khan, R. Stahl, D. Mueller-Gritschneder, and J. Boutellier, "Neural Network-based Vehicle Image Classification for IoT Devices," in IEEE SiPS, 2019.

[37] J. Boutellier, B. Tan and J. Nurmi, "Fault-Tolerant Collaborative Inference through the Edge-PRUNE Framework," In ICML DyNN Workshop, 2022.