Towards open and expandable cognitive AI architectures for large-scale multi-agent human-robot collaborative learning

Georgios Th. Papadopoulos, Member, IEEE, Margherita Antona, and Constantine Stephanidis

Abstract—Learning from Demonstration (LfD) constitutes one of the most robust methodologies for constructing efficient cognitive robotic systems. Despite the large body of research works already reported, current key technological challenges include those of multi-agent learning and long-term autonomy. Towards this direction, a novel cognitive architecture for multi-agent LfD robotic learning is introduced in this paper, targeting to enable the reliable deployment of open, scalable and expandable robotic systems in large-scale and complex environments. In particular, the designed architecture capitalizes on the recent advances in the Artificial Intelligence (AI) (and especially the Deep Learning (DL)) field, by establishing a Federated Learning (FL)-based framework for incarnating a multi-human multi-robot collaborative learning environment. The fundamental conceptualization relies on employing multiple AI-empowered cognitive processes (implementing various robotic tasks) that operate at the edge nodes of a network of robotic platforms, while global AI models (underpinning the aforementioned robotic tasks) are collectively created and shared among the network, by elegantly combining information from a large number of human-robot interaction instances. Pivotal novelties of the designed cognitive architecture include: a) it introduces a new FL-based formalism that extends the conventional LfD learning paradigm to support large-scale multi-agent operational settings, b) it elaborates previous FL-based self-learning robotic schemes so as to incorporate the human in the learning loop, and c) it consolidates the fundamental principles of FL with additional sophisticated AI-enabled learning methodologies for modelling the multi-level inter-dependencies among the robotic tasks. The applicability of the proposed framework is explained using an example of a real-world industrial case study (subject to ongoing research activities) for agile production-based Critical Raw Materials (CRM) recovery.

Index Terms—Learning from demonstration, human-robot interaction, artificial intelligence, federated learning

I. INTRODUCTION

Robot learning through direct interaction with humans constitutes a well-established and highly active research field over the past few decades. In particular, the most dominant learning paradigm in this area, the so-called Learning from Demonstration (LfD), relies on the fundamental principle of robots acquiring new skills by learning to imitate a (human) teacher. LfD (equally termed programming by demonstration, imitation learning, behavioral cloning or apprenticeship learning) concerns multiple aspects of robotics technology, including human-robot interaction, machine learning, machine vision and motor control. Key advantageous characteristics of LfD that have contributed to its widespread adoption and successful application to diverse domains include:

- It enables robot programming by non-expert users;
- It allows time-efficient learning, where task requirements are implicitly learned through demonstrations (and not by explicitly specifying all sequences of robotic actions);
- It enables adaptive robotic behaviors;
- It renders feasible for the robots to operate in complex, unstructured and time-varying environments.

So far, multiple and fundamentally different methods have been investigated and materialized in robotic platforms for capturing, modelling and learning from human feedback in numerous robot manipulation tasks of varying complexity. Depending on the employed human demonstration means, LfD approaches generally fall into three main categories:

- Kinesthetic teaching, which enables the human user to demonstrate by physically moving the robot through the desired motions;
- Teleoperation, which requires an external input to guide the robot through a joystick, graphical user interface or other means;
- Passive observation, where the robot learns from passively observing the user behavior.

A critical aspect in the design and deployment of any LfD scheme concerns the adopted methodology for refining the robot learned policies. In particular, different types of approaches have been introduced:

- Reinforcement learning, where a policy to solve a problem is learned via trial and error;
- Optimization, which targets to find an optimal solution based on given criteria;
- Transfer learning, where knowledge of a task or a domain is used to enhance the learning procedure for another task;
- Apprenticeship learning, where demonstrated samples are used as a template for the desired performance;
- Active learning, where the robotic agent is able to query an expert for the optimal response to a given state and to use these active samples to improve its policy.
• Structured predictions, which is based on the fundamental consideration that an action is regarded as a sequence of dependent predictions [21].

Regarding implementation and deployment aspects of the designed LfD mechanisms, these have been dominated by the adoption of Machine Learning (ML) techniques, whose continuously increasing modelling and representation learning capabilities have correspondingly led to more robust and fine-grained LfD learning potentials. In particular, multiple and diverse ML approaches have been employed, ranging from Gaussian Mixture Models (GMMs) [22], Hidden Markov Models (HMMs) [23] and Dynamic Movement Primitives (DMPs) [24] to, more recently introduced, Recurrent Neural Networks (RNNs) [25] and Convolutional Neural Networks (CNNs) [26], to name a few.

Despite the plethora of works in the LfD field, critical research challenges of paramount importance, namely multi-agent learning [13] and long-term autonomy [27], need to be reliably addressed, in order to promote the wide-spread use of robots in open and complex environments. The latter need to be also investigated in conjunction with typical challenges in Human–Robot Interaction (HRI) schemes, like developing appropriate user interfaces, variance in human performance, variability in knowledge across human subjects, learning from noisy/imprecise human input, learning from very large or very sparse datasets, incremental learning, etc.

In this paper, a novel cognitive architecture for large-scale LfD robotic learning is introduced, targeting to reliably address the currently most critical challenges (and of outstanding importance) in the field, namely those of multi-agent learning, long-term autonomy and deployment of open, scalable and expandable robotic systems. The designed architecture takes advantage of the recent advances in the Artificial Intelligence (AI) (and especially the Deep Learning (DL)) field, by establishing a Federated Learning (FL)-based framework for incarnating a multi-human multi-robot collaborative learning environment. The fundamental conceptualization relies on employing multiple AI-empowered cognitive processes [28] [29] [30] [31] (implementing various robotic tasks, like sensing, navigation, manipulation, control, human-robot interaction, etc.) that operate at the edge nodes of a network of robotic platforms (i.e. adopting a decentralized edge computing setting), while global AI models (underpinning the aforementioned robotic tasks) are collectively created and shared among the network, by elegantly combining information from a large number of human-robot interaction instances. The designed cognitive AI architecture exhibits the following main novel advantageous characteristics, which significantly broaden the current capabilities of LfD learning schemes:

• It extends the current conventional LfD robotic learning paradigm to support large-scale multi-agent operational settings, by introducing a new FL-based formalism, which appropriately amplifies the algorithmic aspects of the conventional FL mechanism;
• It elaborates previous FL-based self-learning robotic schemes, by incorporating the human in the learning loop (using intuitive and informative HRI mechanisms), while including diverse strategies for fine-grained adaptive analysis and integration of multi-human feedback in the Neural Network (NN) parameter update process, namely user weighting, parameter weighting and user clustering;
• It consolidates the fundamental principles of FL with additional sophisticated AI-enabled learning methodologies for modelling the inter-dependencies among the robotic tasks and, hence, further reinforcing the robot knowledge/skill acquisition capabilities, namely transfer-, multitask- and meta-learning techniques.

Overall, the designed cognitive AI architecture essentially introduces a large-scale comprehensive human-robot collective intelligence scheme. An example of a real-world industrial case study (subject to ongoing research activities) for agile production-based Critical Raw Materials (CRM) recovery is investigated for explaining the applicability of the proposed framework.

The remainder of the paper is organized as follows: Section II briefly outlines the fundamental principles of the federated learning paradigm, while Section III discusses prior FL-based works in robotics. Additionally, Section IV details the introduced AI-empowered multi-agent LfD cognitive architecture, whose applicability is demonstrated using a real-world industrial CRM recovery case study in Section V. Moreover, current challenges and proposed future research directions, according to the designed cognitive architecture, are summarized in Section VI. Finally, conclusions are drawn in Section VII.

II. FEDERATED LEARNING PARADIGM

Federated learning (or collaborative learning) is a ML paradigm where multiple computational nodes (e.g. computer clusters, PCs, mobile devices, etc.) collaboratively train a global (AI) model under the supervision of and orchestration by a central process (e.g. server, service provider, High-Performance Computing (HPC) infrastructure, etc.), without exchanging data among the nodes of the network (i.e. maintaining the training data of each node locally, in a decentralized way) [32] [33] [34]. Especially the latter characteristic renders FL a by-definition privacy-aware method, which can reliably mitigate many of the systemic privacy risks and costs resulting from traditional centralized ML [35] [36] [37].

The fundamental mechanism of the FL paradigm is illustrated in Fig. 1. In particular, a global AI model, which is aimed to be collaboratively trained and shared among the network, is initially constructed using proxy data (either offline or at the central node). Then, the model is made available to the network and downloaded by each node. Every node encapsulates a local database that is used to estimate improved updates of the global model parameters (e.g. using conventional Stochastic Gradient Descent (SGD) for the case of NN-based AI modules), without making the local (federated) data available to the network. The computed local parameter updates (denoted $ΔW$ in Fig. 1) are asynchronously transmitted back to the central node (using encrypted communication), where an aggregation mechanism is responsible for combining them (often adopting a simple averaging operator) and
investigating relatively straight-forward implementations of FL schemes in specific robotic tasks have been presented and focus on the following main aspects:

- Autonomous navigation: Liu et al. [40] introduce a reinforcement learning approach, coupled with suitable knowledge fusion and transfer learning algorithms, for autonomous navigation of mobile robots using a cloud environment. Additionally, a reinforcement learning-based real-life collision avoidance system for indoors settings with obstacle objects is presented in [41]. Moreover, an imitation learning framework is proposed for generating guidance models for robots in a self-driving task application in [42]. Furthermore, different approaches have also been investigated for the case of Unmanned Aerial Vehicles (UAVs) path planning [43] [44].
- Simultaneous Localization And Mapping (SLAM): Li et al. [45] present a visual-LiDAR SLAM approach, which supports feature extraction and dynamic vocabulary designation in real-time, while operating on a cloud workstation.
- Motion planning: Bretan et al. [46] incorporate principles of ensemble and reinforcement learning, as well as gradient free optimization, for various robotic tasks, including inverse kinematics, controls and planning.
- Visual perception: Zhou et al. [47] examine a differential privacy protection approach to multiple robotic recognition tasks, while balancing the trade-off between performance and privacy.

Taking into account the above analysis, it can be seen that the literature has so far focused on relatively straight-forward implementations of the FL paradigm and the following key limitations are identified: a) current methods examine only individual robotic tasks in an isolated way, while in typical real-world applications multiple tasks, as well as their cross-correlations, should be simultaneously examined, and b) current approaches are only constrained in self-learning scenarios; however, investigating sophisticated AI-empowered HRI schemes and, hence, involving the human in the learning loop would introduce numerous advantageous characteristics, especially for demanding and fine-grained robotic tasks (e.g. more precise guidance, more time efficient learning, inspection of learning procedure, etc.).

IV. MULTI-AGENT LfD COGNITIVE AI ARCHITECTURE

In this section, a novel open expandable cognitive AI-empowered architecture for multi-agent human-robot collaborative learning is introduced. The ultimate goal is to provide reliable solutions to current critical challenges faced by robotic systems (and especially within the particular LfD field), namely those of multi-agent learning [13] and long-term autonomy [27]; achieving the latter will in turn facilitate the deployment and wide-spread use of robots in open and complex environments. The fundamental conceptualization behind the designed architecture is to leverage the recent technological advancements in the field of AI (focusing mainly on the use of the FL paradigm and closely related technologies) and to transfer them to the robotics application area for extending the
capabilities of the current techniques. The designed cognitive AI architecture, whose high-level representation is illustrated in Fig. 2, incarnates a multi-human multi-robot collaborative learning environment that is composed of the following main reference entities: a) the robotic platform, b) the human, and c) the collective (FL) cognitive AI layer. Detailed analysis of the formalisms and roles of the aforementioned entities are provided in the remaining of the section, while a summary of the main mathematical symbols used is given in Table I.

A. Robotic platform

Under the proposed conceptualization, each robotic platform consists of the following main layers: a) the sensor, b) the robot, and c) the task one. In particular, the sensor layer includes the set of sensing devices that allow the robotic platform to perceive and to collect critical information about the surrounding environment. The set of potentially supported types of sensors, which can be significantly broad and also depends on the particular application case, is defined as follows:

\[ S = \{s_i| i \in [1, I], i \in \mathbb{N}, I \in \mathbb{N}\} \]
\[ = \{\text{Vision, Light, Temperature, Chemical, Force, Acoustic, Gas, Motion, Magnetic, Pressure, Position, ...}\} \quad (1) \]

Regarding the robot layer, this refers to the actual mechatronic equipment to be deployed. Depending on the specific operational scenario, multiple types of robots can be used, supporting different requirements for mobility, positioning, manipulation, communication, size, payload, etc. The set of available types of robots is denoted:

\[ R = \{r_j| j \in [1, J], j \in \mathbb{N}, J \in \mathbb{N}\} \]
\[ = \{\text{Arm, AGV, Humanoid, UAV, Vehicle, Industrial, ...}\} \quad (2) \]

With respect to the task layer, this concerns the types of policies and activities that the robotic platform will be required to implement. These may cover a large set of possible perception, cognition, motor and interaction functionalities, which will inevitably need to be appropriately adapted, apart from the specific application requirements at hand, also to the particularities of the employed robot type \( r_j \). The set of potential types of robot tasks is defined as follows:

\[ T = \{t_k| k \in [1, K], k \in \mathbb{N}, K \in \mathbb{N}\} \]
\[ = \{\text{Sensing, Navigation, Manipulation, Control, Human – robot interaction, ...}\} \quad (3) \]

Taking into account the above-mentioned formalisms, a robotic platform \( P_l \) can be fully specified as follows:

\[ P_l = \{S_l, R_l, T_l|S_l \subseteq S, R_l \subseteq R, T_l \subseteq T\}, l \in [1, L], \quad (4) \]

where \( L \) denotes the total number of robotic platforms present in the examined cognitive environment.

B. Human

Within the designed multi-agent collaborative learning environment, the human factor constitutes a fundamental building block for simultaneously: a) transferring fine-grained and sophisticated skills to the robot and b) supervising/inspecting the learned robot behaviors (i.e. in principle guiding the overall robotic learning process). Although all types of LfD methodologies (namely kinesthetic teaching, teleoperation and passive observation, as detailed in Section I) are supported by the introduced cognitive AI architecture, the adoption of a teleoperation-based approach is considered to exhibit significant advantageous characteristics. In particular, an intuitive and sophisticated teleoperation scheme is foreseen that is based on the combined used of Augmented Reality (AR) visualization mechanisms and eXplainable AI (XAI) technologies.

Regarding AR techniques, they have so far been shown to be beneficial for enabling the human operators to be simultaneously aware of the actual processes that take place in the physical environment (e.g. a factory) and to be continuously updated with valuable information related to the underlying automatic control procedures. In the current cognitive architecture, AR tools are employed in order to allow the human user to perform a physical inspection of the robot exhibited behaviors (and hence to identify malfunctions, hazardous situations, deviations from desired policies, etc.), while at the same time being constantly provided with key detailed insights about the AI processes being applied (e.g. the AI modules being used, their estimated outputs, how specific decisions are reached, etc.). To this end, this AR-grounded setting constitutes an efficient and user-friendly way to simultaneously examine the convergence of the physical (robot) and AI (software) worlds. Additionally, while also of paramount importance, the designed AR setting makes two key functionalities available to the human user: a) to provide feedback regarding the possible corrections in the exhibited robot behaviors (when a deviation from the desired targets is observed) and b) to receive full control of the robot performed actions, through a teleoperation-based robot policy definition scheme. Under certain circumstances (e.g. involvement of very large-scale application settings, like industrial
operating plants), Virtual Reality (VR) technologies can also be used in conjunction (e.g. virtual factory) for overall process monitoring. It needs to be mentioned that the set of human operators involved in the designed cognitive architecture is denoted $H = \{h_m | m \in [1,M], m \in \mathbb{N}, M \in \mathbb{N}\}$.

Concerning XAI methods, their fundamental aim is to improve trust and transparency of AI-based systems, by attempting to provide valuable insights or to directly explain the decision making process of AI procedures [48]. Under the current conceptualization, XAI techniques are adopted in order to provide precise explanations/insights to the human operator regarding the deviation of the robot behavior from the desired one, i.e. enabling an in depth inspection of the robot behavior. The XAI-based generated explanations are provided to the human user through the aforementioned AR/VR visualization interfaces. Consequently, the human user is capable of providing more accurate guidance to the robot (through means of provision of feedback or teleoperation, as discussed above) and, hence, to supervise the overall robot learning process more closely. Depending on the particular type of sensor $s_i$ and task $t_k$, different model-specific or model-agnostic XAI methods can be employed [49].

C. Collective (FL) cognitive AI layer

The introduced collective (FL) cognitive AI layer constitutes the core entity in the designed architecture that is responsible for creating, updating and distributing multiple AI modules, which underpin the various robotic processes implemented in the examined application setting. These AI modules are collectively created and maintained by simultaneously aggregating knowledge/feedback from a large-scale human-robot interaction set (adopting the fundamental mechanisms of the LfD methodology), while aiming to address the current challenges of multi-agent learning and long-term autonomy. For achieving the latter goals, the designed open and expandable cognitive AI architecture is grounded on principles of the FL approach, whose fundamental mechanism is explained in Section III. The main building blocks of the introduced cognitive AI layer, their formalism and detailed explanation of their functionalities, is provided in the followings:

Fig. 2. High-level representation of the introduced multi-agent cognitive AI collaborative learning architecture.
Network nodes: Every robotic platform \( P_i \) corresponds to a network node of the defined architecture, in accordance to the fundamental FL mechanism illustrated in Fig. 1. Each \( P_i \) stores locally the generated data, which in principle contain information collected from the set of sensors \( S_i \) incorporated by \( P_i \). It needs to be highlighted that feedback information can be obtained by any \( P_i \) from the interaction with any human subject \( h_m \) present in the application environment, i.e. it is considered that any teacher \( h_m \) can inspect and provide guidance to any performing robotic platform \( P_i \).

AI models: The designed cognitive architecture aims to address the needs related to the deployment of large-scale AI-driven human-robot environments, simultaneously supporting multiple combinations of sensor \( s_i \), robot \( r_j \) and task \( t_k \) types. For achieving that, a broad set of individual AI models denoted \( W_β \) $\leftarrow f_1(S_β, R_β, T_β) $, underpinning the various implemented cognitive processes present in the examined environment, is considered, where \( S_β \subseteq S, R_β \subseteq R, T_β \subseteq T, β \in [1, B], β \in \mathbb{N}, B \in \mathbb{N} \) and \( f_1(.) \) implies a generalized function or process that defines the exact NN-based materialization of model \( W_β \) while considering \( S_β, R_β \) and \( T_β \) as input parameters. The overall goal of the introduced FL-grounded environment is to construct global \( W_β \) models (Section 1), by exploiting data from a large-scale human-robot interaction set of distributed sources, while each network node can maintain a local/customized version of \( W_β \).

Learning methodology: Regarding the specific methodology to be followed for refining the robot learned policies (i.e. for updating the AI models \( W_β \)), different options can be investigated (e.g. reinforcement learning, transfer learning, active learning, etc.), as discussed in Section 1. The most suitable selection depends on the particularities of the application domain and each individually examined \( W_β \).

Local parameter updates: Regardless of the particular learning methodology selected, each robotic platform \( P_i \) can estimate updates for any AI model \( W_β \) that is associated with, using its locally stored data. More specifically, the following local parameter update mechanism is applied:

\[
W_β^{l,m} \leftarrow W_β^{l,m} - lr \cdot \nabla L_β(W_β^{l,m}, δ^l),
\]

where \( W_β^{l,m} \) is the local/customized version of the global model \( W_β \) with respect to human teacher \( h_m \), \( lr \) is the local learning rate, \( \nabla \) denotes the gradient of a function, \( L_β(.) \) represents the loss function defined for \( W_β \) and \( δ^l \) is the locally stored dataset. Consequently, the local parameter updates, which are iteratively estimated, to be sent to the central node (aggregator) are computed as follows:

\[
\Delta W_β^{l,m} = W_β^{l,m} - W_β
\]

From the above formalization, it can be seen that the designed architecture allows local parameter updates to be estimated separately for each human teacher \( h_m \), i.e. also leading to ‘personalized’ versions of each model \( W_β \).

User profiling: Incorporating human feedback constitutes a fundamental part of the LiD approach and, consequently, of the current cognitive architecture. However, the latter poses additional challenges to the problem formulation that need to be efficiently addressed (e.g. variance in human performance, variability in knowledge across human subjects, learning from noisy/imprecise human input, learning from very large or very sparse data sets, incremental learning, etc.). Towards this direction and in parallel with the \( W_β^{l,m} \) estimation process, an individual user profile \( Q_β^{l,m} ← f_2(δ^l, S_m) \) is constructed for every human subject \( h_m \) at every node \( P_i \), given the appropriate sensorial data \( S_m \) to model the observed human behavior and a generalized function \( f_2(.) \) that defines the exact (NN-based) implementation of the user profile while considering \( S_m \) as input parameters. The aim of \( Q_β^{l,m} \) is to cover physical (e.g. human actions, physical capabilities, etc.), cognitive (e.g. intention, personality, etc.) and social (e.g. non-verbal cues, emotions, etc.) aspects, in order to model and efficiently interpret the exhibited human activity [50]. Under the current conceptualization, DL-based approaches are considered for generating \( Q_β^{l,m} \), as in [51] and [52], aiming at combining increased modelling capabilities and easier integration to the designed FL-based framework.

Global model update: Having computed the local parameter updates \( \Delta W_β^{l,m} \) and the corresponding user profiles \( Q_β^{l,m} \) (using locally generated and processed data \( δ^l \)), these are sent to the central node (aggregator) so as to periodically produce updated versions of the global \( W_β \) models. According to the conventional FL mechanism, an updated version \( W_β \) of each individual \( W_β \) is generated on a regular basis, by applying a simple average operator, as follows:

\[
W_β = W_β + lr^g \cdot Γ_β
\]

\[
Γ_β = \frac{1}{Λ_β} \sum_{l,m} \Delta W_β^{l,m}, \quad (7)
\]

where \( lr^g \) denotes the global learning rate and \( Λ_β \) the total number of received \( \Delta W_β^{l,m} \) updates. The strength of the above-mentioned mechanism lies on incorporating a very large number \( Λ_β \) of samples and multiple iterative updates of \( W_β \) that will likely lead to the convergence to well-performing and robust \( W_β \) models, while the network nodes \( P_i \) contributing in (7) may be sampled out of the available ones (usually in a random way). However, combining information (\( \Delta W_β^{l,m} \)) related to different human subjects \( h_m \) (that presumably exhibit highly diverse and varying behavior) is in turn very likely to lead the FL process to be confined to a local maximum or even to a non-convergence of the FL procedure; hence, jeopardising the overall learning process.

1) Incorporation of multi-user feedback information: For robustly confronting the observed variance in the behavior of the large number of involved human teachers \( h_m \), the designed cognitive architecture (apart from possible sensorial data \( s_i \) pre-processing for invariance incorporation) encompasses the estimated user profiles \( Q_β^{l,m} \) in the global model update process, modifying (7) to the following general formalism:

\[
Γ_β = \Phi(\{Q_β^{l,m}\}, \{\Delta W_β^{l,m}\}), \quad (8)
\]

where \( Φ(.) \) denotes a generalized function that combines the available \( Q_β^{l,m} \) and \( \Delta W_β^{l,m} \), estimated at every network node \( P_i \). Depending on the particularities of the selected application domain (e.g. supported \( W_β, S, R, T, \) etc.), the following
main materializations of $\Phi(\cdot)$, while also being possible to be combined, are considered:

- **User weighting**: Under this conceptualization, the contribution of each human teacher $h_{m}$ is modulated by a different weight factor based on his/her exhibited behavior, as follows:

$$\Gamma_{\beta} = \frac{1}{E_{\beta}} \sum_{(l,m)} \varepsilon(Q^{l,m}) \cdot \Delta W_{\beta}^{l,m}$$

$$\varepsilon(Q^{l,m}) = \frac{1}{\|\Phi(h_{m}) - Q^{l,m}\|}$$

$$E_{\beta} = \sum_{(l,m)} \varepsilon(Q^{l,m}), \quad (9)$$

where $\varepsilon(Q^{l,m})$ denotes the weight factor for each $h_{m}$ at every $P_{l}$, $Q^{h_{m}}$ the corresponding global user model (e.g. estimated through the same FL-based mechanism used for constructing $W_{\beta}$) and $\|\|$ a similarity score metric (e.g. Euclidean distance between the parameters of the involved user models).

- **Parameter weighting**: The fundamental consideration lies on performing sensitivity analysis $\Phi(\cdot)$ for estimating the degree of correlation among the parameters of $W_{\beta}$ and $Q^{h_{m}}$, i.e. emphasizing on how the exhibited behavior of each user $h_{m}$ affects individual aspects/parameters of $W_{\beta}$, according to the following formalism:

$$\Gamma_{\beta} = \frac{1}{R_{\beta}} \sum_{(l,m)} r_{\beta}^{l,m} \cdot \Delta W_{\beta}^{l,m}$$

$$r_{\beta}^{l,m} = U(Q^{l,m}, W_{\beta}^{l,m})$$

$$R_{\beta} = \sum_{(l,m)} r_{\beta}^{l,m}, \quad (10)$$

where function $U(\cdot)$ realizes sensitivity analysis for estimating the impact of parameters $Q^{l,m}$ on the respective ones of $W_{\beta}$, while matrix $r_{\beta}^{l,m}$ summarizes the estimated correlations.

- **User clustering**: The main principle behind this approach, the so called ‘Multi-Center Federated Learning’ $\Phi(\cdot)$, considers the generation of multiple instances $W_{\beta_{n}}^{l,m}$ of each global model $W_{\beta}$, in order to better capture the heterogeneity of data distributions across different users. In particular, each human teacher $h_{m}$ at each local network node $P_{l}$ is associated with a single $W_{\beta_{n}}$, the latter is iteratively updated, by elaborating $\Phi(\cdot)$, as follows:

$$\Gamma_{\beta} = \frac{\sum_{(l,m)} 1}{\sum_{(l,m)} \theta_{\beta_{n}}^{l,m} \Delta W_{\beta}^{l,m}}, \quad (11)$$

where $\theta_{\beta_{n}}^{l,m} = 1$, if teacher $h_{m}$ at node $P_{l}$ is associated with $W_{\beta_{n}}$, and $\theta_{\beta_{n}}^{l,m} = 0$, otherwise.

It needs to be highlighted that all above-mentioned variants of function $\Phi(\cdot)$ in (9)-(11) are considered to be implemented in a neural-network form; hence, rendering the overall approach end-to-end learnable within the same integrated FL scheme.

2) **Incorporation of cross-task correlation information**: Complementary to the integration of information from multiple human teachers $h_{m}$ (as detailed in Section [V-C1]), the designed cognitive layer puts also particular emphasis on analysing, modelling and exploiting the correlations among the multiple (and often co-occurring) robotic tasks, which are controlled by the AI models $W_{\beta}$. Towards this direction, the aforementioned FL-based mechanisms for incorporating multi-human feedback information is further elaborated and enhanced, by integrating the following sophisticated AI-empowered learning capabilities:

- **Transfer learning**: Federated transfer learning constitutes a suitable methodology for addressing cases where different AI-driven processes share an overlap in the respective feature space, aiming at exploiting the underlying data correlations and building models collaboratively $\Phi(\cdot)$ [56] [57]. In particular, the conventional FL mechanism (described in Section [V-C1]) considers the separate construction of each global $W_{\beta}$ model, using the locally generated data $\delta_{l}$, by adopting the following general type of loss function during the training step:

$$L_{\gamma} = \sum_{\beta, (l,m)} L(\hat{Y}(W_{\beta}), Y(W_{\beta})) \quad (12)$$

where $L_{\gamma}$ denotes the employed global loss function, $L_{\beta}$ corresponds to the loss term with respect to each individual $W_{\beta}$ and $Y(\cdot)$ is the estimated, targeted (ground truth) output of $W_{\beta}$, respectively. In order to achieve feature transfer learning, the following alignment loss factor is added to $L_{\gamma}$ in (12):

$$L_{2} = \sum_{(\beta_{1,\beta_{2}}, (l,m)} V(W_{\beta_{1}}, W_{\beta_{2}}) \quad (13)$$

where $V(\cdot)$ denotes an alignment/similarity measure between $W_{\beta_{1}}$ and $W_{\beta_{2}}$ (e.g. Euclidean distance $||\cdot||$ between the model parameters, as in (9)). It needs to be highlighted that the above mentioned transfer learning mechanism is applicable for AI models that exhibit similar patterns and correlations in the underlying data space, which inevitably implies overlaps and similarities between the respective $S_{\beta_{1}}$ and $S_{\beta_{2}}$ sets.

- **Multi-task learning**: According to the formalisms proposed in the literature so far for federated multi-task learning, the fundamental aim is posed as simultaneously constructing separate, but related, AI models at each network node [58] [59]. The latter requires, among others, the formulation of a so called precision matrix that encodes the inter-relations among the models, which can be either explicitly defined a priori or learned directly from the data. Under the current conceptualization, the problem of federated multi-task learning is re-formulated so as to allow the simultaneous learning of multiple global models $W_{\beta}$ that correspond to related (and often co-occurring) robotic tasks. The respective loss function
to be used during training has the following general form:

$$\mathcal{L}_g = \sum_{\beta}(\bar{Y}(W_\beta), Y(W_\beta)) + \mathcal{X}(A, \Omega)$$

$$A = [W_1 \ W_2 \ ... \ W_B], \quad (14)$$

where $A$ is a weight matrix produced by the concatenation of the individual $W_\beta$ model parameters and

$$\Omega \in \mathbb{R}^{B \times B}$$

is the so called precision matrix. Function $\mathcal{X}()$ summarizes the defined assumptions of the federated multi-task learning problem, where a bi-convex formulation is often selected, as follows:

$$\mathcal{X}(A, \Omega) = \frac{\lambda}{2}tr(A\Omega A^T)$$

$$\Omega^{-1} \geq 0, \quad tr(\Omega^{-1}) = 1, \quad (15)$$

where $\lambda$ is a constant and $tr(.)$ denotes the trace of a matrix.

- **Meta learning**: The principal goal of the FL mechanism, as detailed in the beginning of Section [V.C] is to collectively create robust and powerful global AI models $W_\beta$ (14), by concatenating model updates $\Delta W_{\beta}^{l,m}$ (7) from multiple human teachers $h_m$ associated with the various network nodes $P_l$. However, there are application cases where the fundamental aim is not (only) to optimize the performance of the global models $W_\beta$, but rather to maximize the efficiency of the local ones $W_{\beta}^{l,m}$ (5), e.g. when a new human teacher $h_m$ is introduced to the cognitive environment (and an accurate initialization of the respective $W_{\beta}^{l,m}$ models is required) or when the personalization ability of the overall system is critical (for example when developing customized human-robot interaction schemes). For addressing the latter requirements, the designed cognitive architecture incorporates means of so called meta learning techniques. Towards this direction, different meta learning methodologies have been proposed in the literature with Model Agnostic Meta Learning (MAML) receiving particular attention. The latter inevitably poses significant challenges towards their efficiency management, which in turn requires an overall continuous adjustment of the recycling plant work-plan and internal functionality. It needs to be highlighted that multiple types of waste streams (that contain valuable CRMs) are considered and processed in parallel, as mentioned in the beginning of Section [V.A] Upon their receipt (and after unpacking, if ongoing research activities) in the domain of Critical Raw Materials (CRM) recovery. Robotic platforms constitute a particularly suitable solution for the selected field, since they can undertake and automate numerous laborious, repetitive, tedious, stressful, harmful and hazardous human worker tasks, especially in the early stages of the recycling process (e.g. object dismantling, housing removal, component extraction, etc.). Overall, the CRM recovery pipeline inevitably needs to undergo an agile production operational methodology, since a) every recycling plant typically supports multiple waste streams of different type and nature, where the input materials arrive in an unsettled way and of significantly varying quantity, quality and composition, and b) the demands for the output recycled materials (posed by secondary market stakeholders, further recycling operators, etc.) also change (often rapidly) over time. The latter essentially poses the critical requirement for a high degree of adaptability to the involved robotic platforms, accompanied by the increased need for reinforced long-term autonomy (since new types of CRM materials and needed manipulations/operations are constantly encountered). To this end, the introduced cognitive AI architecture for multi-agent LfD learning, operating complementarily to a centralized factory-level orchestration methodology for agile production, constitutes an elegant choice for enabling the robotic platforms to continuously acquire new skills (from human demonstration). A high-level functional diagram of the envisaged agile production CRM recovery plant is illustrated in Fig. [3].

A. Roboticized processing steps

Throughout the overall recycling plant operational pipeline, the input waste materials (e.g. laptops, personal computers, smartphones, tablets, TVs, batteries, etc.) undergo subsequent processing and manipulation steps, targeting to extract and group individual constituent device components (e.g. casings, plastics, capacitors, coolers, Printed Circuit Boards (PCBs), etc.) with homogeneous composition in terms of integrant critical raw materials (e.g. cobalt, lithium, phosphorus, magnesia, bauxite, etc.). For realizing the latter, different categories of robotic platforms with varying types of assigned tasks need to be deployed in each of the recycling plant’s main processing steps (namely material routing, device dismantling and component sorting), as detailed in the followings.

1) **Material routing**: The first step in the envisaged recycling plant work-cycle concerns how the different types of waste materials are being introduced to the processing pipeline. Under the current conceptualization, different means of plant input are foreseen (e.g. aggregated piles, received packages, delivered containers, etc.). The critical characteristic is that the input materials arrive in an unsettled way and in significantly varying pace, quantity, quality and composition. The latter inevitably poses significant challenges towards their efficient management, which in turn requires an overall continuous adjustment of the recycling plant work-plan and internal functionality. It needs to be highlighted that multiple types of waste streams (contain valuable CRMs) are considered and processed in parallel, as mentioned in the beginning of Section [V.A]. Upon their receipt (and after unpacking, if
needed), the different waste materials are identified/classified and assigned/introduced to a waste stream processing line. This step involves the transfer of the input waste materials to a dynamically assigned processing line of the factory. This dynamic ‘material routing’ process involves the use of Automatic Guided Vehicles (AGVs), each mounted with a robotic arm (to implement the necessary pick and place actions) and a conveyor belt mechanism (for further facilitating the loading/unloading process). It needs to be highlighted that a varying number of processing lines for each individual waste stream type are also adaptively defined, according to the incoming waste status as well as the targeted overall plant output at each time instant.

2) Device dismantling: Having performed the initial waste stream classification and its assignment to a processing line, the core building block of the overall solution is applied for realizing the ‘device disassembly/dismantling’ step. In particular, each processing line comprises a conveyor belt, where the pace and speed of operation are dynamically controlled, taking into account the materials present or being processed. Along both sides of each conveyor belt, multiple workspaces are installed. The latter constitute the physical locations where the fine-grained manipulation of the input devices/materials (e.g. dismantling task) occurs. At each workspace, a multi-robot cell (consisting of multiple identical robotic arms) is dynamically assigned, based on run-time material recovery needs. The total number of multi-robot cells operating in each line is also adaptively and centrally defined, based on overall factory input/output targets and operational status. Each of the involved robot cells is equipped with a suitable tool changer, so as to support the use of multiple end-effectors and, hence, to enable different and diverse fine-grained manipulation tasks (e.g. cutting, unscrewing, drilling, breaking, etc.). The output of this step is a set of extracted and sorted constituent components of interest for every manipulated device, placed in appropriate collection baskets.

3) Component sorting: After the input waste materials are processed and the components of interest are extracted, the so called ‘component sorting’ step is implemented, where AGVs are responsible for transferring the components from the disassembly line to the deposit location of the plant (e.g. packaging stations, material separation machines, etc.). The AGVs are equipped with effective basket mounting/unmounting and advanced SLAM navigation capabilities, in order to smoothly and safely operate in complex, time-varying environments and in the presence of humans.

B. Integration of the cognitive AI architecture

In order to realize the above mentioned complex waste management activities in an autonomous, productive, efficient and safe way, each employed robotic platform $P_i$ is equipped with sophisticated edge computing AI-empowered cognitive
technologies that reinforce its operational capabilities (e.g. working environment registration, human behaviour analysis, waste manipulation, motion planning, navigation, human-robot interaction, safety control, etc.); the latter are implemented in the form of respective sophisticated AI models $W_\beta$. The most critical part for ensuring the long-term autonomy of the introduced system concerns the point where the robotic platforms $P_l$ and the human operators $h_m$ interact on the factory floor, targeting in principle the bootstrapping of the robotic activities as well as the acquisition of new skills from the side of the robotic agents. In particular, the human workers operate in a supervisory and proactive way, while they are equipped with AR technologies in order to receive real-time and accurate insights on the processes performed at the factory- and each individual workspace-level. The latter insights include detailed information on the status of the specific tasks performed by the robotic platforms (i.e. which steps have been implemented and which are planned to be performed) and regarding the robotic cognitive/inference operations (e.g. which objects have been recognized by the robot, how a given decision has been reached, what type of motion planning policies have been estimated, etc.). Especially for the latter case, the use of XAI tools enables the interpretation of the exhibited robot behaviour, i.e. the identification of the root causes/procedures that led an AI-driven robotic platform to take specific decisions/actions. In this context, whenever a robot fails to complete a task, there is a high degree of uncertainty or the human worker identifies a deviation and decides to intervene, the robot pauses its operations and waits for human feedback. Subsequently, through the use of user-friendly and efficient AR technologies, the human worker is initially informed of the robot operational status (as described above), identifies the root cause of the possible malfunction and through the appropriate AR-based interaction/communication means guides the robot on how to successfully elaborate the task at hand. Apart from the in situ tuning phase, the principal goal of the overall system is for the robotic platforms to adaptively learn and in the long-term adapt their AI cognitive models $W_\beta$, based on the feedback received by the human workers (and modulated by the estimated user profiles $Q^{l,m}$). For incarnating this multi-human multi-robot collective intelligence vision, the introduced FL-based multi-agent LfD cognitive AI architecture (as detailed in Section IV) is applied. It needs to be reminded that the designed FL-based scheme is by definition privacy-aware, since no data, apart from NN parameter updates, captured at each robot/node location $P_l$ are sent to the network.

VI. CHALLENGES AND FUTURE RESEARCH DIRECTIONS

The current work focuses on introducing a novel cognitive AI architecture, targeting to provide reliable solutions to current obstacles in (large-scale) LfD robot learning, namely that of long-term autonomy and multi-agent systems. For achieving this, a new FL-based AI formalism is introduced for enabling open and expandable human-robot collaborative learning schemes. The designed architecture, which significantly broadens the capabilities of previous FL-based and LfD robotic systems, suggests specific and concrete research directions for each of the challenges met (also in relation to the respective state-of-art solutions), as detailed in Section V. In this section, the key elements of the designed environment that are crucial to its success, which at the same time constitute both critical technological challenges and promising research directions in the fields of LfD, AI and HRI, are briefly summarized below:

- **Incorporation of multi-user feedback information:** Although numerous LfD approaches have been introduced in the past, multi-agent learning remains as one of the most challenging open issues in the field [13]. The latter comprises the central goal of the introduced cognitive AI architecture, which relies on the incorporation of FL-based technologies reinforced with a generalized function $\Phi(\cdot)$ (defined in (8)) for specifying how to combine information (i.e. model updates $\Delta W^{l,m}_\beta$) from multiple human subjects $h_m$ (making use of the respective estimated profiles $Q^{l,m}$). Three variants of function $\Phi(\cdot)$ are introduced (Section IV-C1), namely user weighting, parameter weighting and user clustering. The specific choice to be made depends on the particularities of the selected application domain (e.g. the number of involved human workers, the number and types of supported AI models, the type and number of employed sensors, etc.).
- **Incorporation of cross-task correlation information:** LfD methods have so far focused on introducing invariance to the observed human behavior for a given task. However, the designed architecture puts emphasis on exploiting the correlations among different tasks (i.e. among the various $W_\beta$ or $W^{l,m}_\beta$ models); hence, significantly elaborating current LfD principles. Three individual methodologies are proposed towards this direction (Section IV-C2), namely transfer, multi-task and meta learning. The selection of the specific approach to be used (or a combination of them), depends again on the individual characteristics and goals of the targeted application (e.g. number and type of $W_\beta$ models involved, if new human users are expected to be involved and how often, etc.).
- **DL-empowered user profiling:** There has been an extensive body of research activity for creating robust user profiles, using various ML techniques, over the recent years. However, further focusing on implementing such models following the DL paradigm would likely lead to significant performance gains and increased robustness, while, importantly, the use of NNs [51] for building $Q^{l,m}$ models (8) would enable the incorporation of fully end-to-end trainable systems and $Q^{l,m}$ to be constructed through the designed FL-based cognitive environment. Further challenges comprise the use of multi-modal information from multiple non-invasive sensorial devices, while simultaneously modelling the variance in the exhibited human behaviour (among the same or different individuals).
- **Deployment of FL technologies in robotic systems:** FL itself often poses significant deployment challenges of various types (e.g. low connectivity, increased number
of network nodes, latency in communication, etc.) \[38\] \[65\]. In the context of a large-scale industrial robotic setting, such challenges will have an increased importance, while the expected levels of system robustness and interoperability will inevitably need to meet higher industrial standards as well. To this end, reliable FL-based solutions would likely need to capitalize on, apart from algorithmic optimizations of the FL mechanism, the capabilities provided by additional emerging technologies (e.g., 5G/6G network connectivity, quantum computing, hardware AI implementations, etc.).

- **Human-centred eXplainable AI**: Over the recent years, an extensive body of research has been devoted on investigating various XAI methodologies, resulting in numerous diverse approaches and promising results \[49\]. However, the focus has so far been placed on addressing the ‘explainability’ aspect (i.e., to identify the underlying reasons for the exhibited behavior of the AI models), leaving the ‘interpretable’ perspective (i.e., the human users actually understanding the observed AI behavior) largely under-explored \[43\]. The latter becomes even more demanding for cases of non-IT human experts being involved. A promising direction would naturally require the incorporation of Human-Computer Interaction (HCI) and human sciences principles, e.g., through the use of dynamic visualizations, question-answering schemes, interactive mechanisms, etc.

- **Addressing of (cyber-)security, personal data and privacy/ethics issues**: Despite the fact that FL is a by-definition privacy aware approach that requires no exchange of actual data (only AI model parameter updates are transmitted) among the network nodes, significant research efforts have been devoted recently towards addressing possible security and privacy preserving gaps (e.g., differential \[35\], model-poisoning \[66\], white-box inference \[67\] attacks, etc.). The latter are often combined with innovative techniques or emerging technologies, such as differential privacy \[68\], homomorphic encryption \[69\], blockchain \[70\], etc. However, explicitly modeling and integrating human user feedback information, through the creation of user profiles \(Q(\cdot,\cdot)\) and their incorporation in the FL mechanism (Section IV-C), inevitably requires the elaboration and extension of the aforementioned methodologies.

### VII. CONCLUSIONS

In this paper, the problem of robot Learning from Demonstration (LfD) was thoroughly investigated and a novel cognitive architecture for large-scale robotic learning was introduced for enabling the robust deployment of open, scalable and expandable robotic system in large-scale and complex environments. The fundamental conceptualization of the designed architecture is grounded on the establishment of a Federated Learning (FL)-based framework for implementing a multi-human multi-robot collaborative learning environment. Pivotal novelties of the designed cognitive architecture that significantly broaden the capabilities of current LfD robotic learning schemes include: a) it introduces a new FL-based formalism that extends the conventional LfD learning paradigm to support large-scale multi-agent operational settings, b) it elaborates previous FL-based self-learning robotic schemes so as to incorporate the human in the learning loop, and c) it consolidates the fundamental principles of FL with additional sophisticated AI-enabled learning methodologies for modelling the multi-level inter-dependencies among the robotic tasks. The applicability of the designed framework was explained through an example of a real-world industrial case study for agile production-based Critical Raw Materials (CRM) recovery. Moreover, detailed analysis of the current technological challenges and future research directions were discussed.

### ACKNOWLEDGMENTS

The work presented in this paper was supported by the ICS-FORTH internal RTD Programme ‘Ambient Intelligence and Smart Environments’.

### REFERENCES

[1] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, “Recent advances in robot learning from demonstration,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, 2020.

[2] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, “A survey of robot learning from demonstration,” *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.

[3] D. Furuta, K. Kutsuzawa, S. Sakaino, and T. Tsuji, “Motion planning with success judgement model based on learning from demonstration,” *IEEE Access*, vol. 8, pp. 73 142–73 150, 2020.

[4] I. A. Seleem, H. El-Hussieny, S. F. Assal, and H. Ishii, “Development and stability analysis of an imitation learning-based pose planning approach for multi-section continuum robot,” *IEEE Access*, 2020.

[5] A. G. Billard, S. Calinon, and R. Dillmann, “Learning from humans,” in *Springer handbook of robotics*. Springer, 2016, pp. 1995–2014.

[6] Z. Zhu and H. Hu, “Robot learning from demonstration in robotic assembly: A survey,” *Robotics*, vol. 7, no. 2, p. 17, 2018.

[7] R. Caccavale, M. Saveriano, A. Finzi, and D. Lee, “Kinesthetic teaching and attentional supervision of structured tasks in human–robot interaction,” *Autonomous Robots*, vol. 43, no. 6, pp. 1291–1307, 2019.

[8] M. Capurso, M. M. G. Ardakani, R. Johansson, A. Robertson, and P. Rocco, “Sensorless kinesthetic teaching of robotic manipulators assisted by observer-based force control,” in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017, pp. 945–950.

[9] T. Zhang, Z. McCarthy, O. Jow, D. Lee, X. Chen, K. Goldberg, and P. Abbeel, “Deep imitation learning for complex manipulation tasks from virtual reality teleoperation,” in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–8.

[10] Y. Xu, C. Yang, J. Zhong, N. Wang, and L. Zhao, “Robot teaching by teleoperation based on visual interaction and extreme learning machine,” *Neurocomputing*, vol. 275, pp. 2093–2103, 2018.

[11] Y. Liu, A. Gupta, P. Abbeel, and S. Levine, “Imitation from observation: Learning to imitate behaviors from raw video via context translation,” in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1118–1125.

[12] C. Finn, T. Yu, T. Zhang, P. Abbeel, and S. Levine, “One-shot visual imitation learning via meta-learning,” *arXiv preprint arXiv:1709.04905*, 2017.

[13] A. Hussein, M. M. Gaber, E. Elyan, and C. Jayne, “Imitation learning: A survey of learning methods,” *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–35, 2017.

[14] M. Zhang, Z. McCarthy, C. Finn, S. Levine, and P. Abbeel, “Learning deep neural network policies with continuous memory states,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 520–527.

[15] T. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al., “Mastering the game of go with deep neural networks and tree search,” *nature*, vol. 529, no. 7587, pp. 484–489, 2016.
[61] Y. Jiang, J. Konečný, K. Rush, and S. Kannan, “Improving federated learning personalization via model agnostic meta learning,” *arXiv preprint arXiv:1909.12488*, 2019.

[62] A. Fallah, A. Mokhtari, and A. Ozdaglar, “Personalized federated learning: A meta-learning approach,” *arXiv preprint arXiv:2002.07948*, 2020.

[63] F. Chen, M. Luo, Z. Dong, Z. Li, and X. He, “Federated meta-learning with fast convergence and efficient communication,” *arXiv preprint arXiv:1802.07876*, 2018.

[64] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, “Federated learning: Strategies for improving communication efficiency,” *arXiv preprint arXiv:1610.05492*, 2016.

[65] Y. Yan, S. Liu, Y. Jin, Z. Qun, S. Zhang, and S. Lu, “Risk minimization against transmission failures of federated learning in mobile edge networks,” *IEEE Access*, 2020.

[66] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, “How to backdoor federated learning,” in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2020, pp. 2938–2948.

[67] M. Nasr, R. Shokri, and A. Houmansadr, “Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning,” in *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2019, pp. 739–753.

[68] S. Kim, “Incentive design and differential privacy based federated learning: A mechanism design perspective,” *IEEE Access*, vol. 8, pp. 187317–187325, 2020.

[69] L. Song, C. Ma, G. Zhang, and Y. Zhang, “Privacy-preserving unsupervised domain adaptation in federated setting,” *IEEE Access*, vol. 8, pp. 143233–143240, 2020.

[70] M. A. Rahman, M. S. Hossain, M. S. Islam, N. A. Alrajeh, and G. Muhammad, “Secure and provenance enhanced internet of health things framework: A blockchain managed federated learning approach,” *IEEE Access*, vol. 8, pp. 205071–205087, 2020.