

Original Paper

# Balancing Between Privacy and Utility for Affect Recognition Using Multitask Learning in Differential Privacy–Added Federated Learning Settings: Quantitative Study

Mohamed Benouis, PhD; Elisabeth Andre, PhD; Yekta Said Can, PhD

Faculty of Applied Computer Science, Augsburg University, Augsburg, Germany

**Corresponding Author:**

Yekta Said Can, PhD  
Faculty of Applied Computer Science  
Augsburg University  
Universitätsstraße 6a  
Augsburg, 86159  
Germany  
Phone: 49 015221351388  
Email: [yekta.can@uni-a.de](mailto:yekta.can@uni-a.de)

## Abstract

**Background:** The rise of wearable sensors marks a significant development in the era of affective computing. Their popularity is continuously increasing, and they have the potential to improve our understanding of human stress. A fundamental aspect within this domain is the ability to recognize perceived stress through these unobtrusive devices.

**Objective:** This study aims to enhance the performance of emotion recognition using multitask learning (MTL), a technique extensively explored across various machine learning tasks, including affective computing. By leveraging the shared information among related tasks, we seek to augment the accuracy of emotion recognition while confronting the privacy threats inherent in the physiological data captured by these sensors.

**Methods:** To address the privacy concerns associated with the sensitive data collected by wearable sensors, we proposed a novel framework that integrates differential privacy and federated learning approaches with MTL. This framework was designed to efficiently identify mental stress while preserving private identity information. Through this approach, we aimed to enhance the performance of emotion recognition tasks while preserving user privacy.

**Results:** Comprehensive evaluations of our framework were conducted using 2 prominent public datasets. The results demonstrate a significant improvement in emotion recognition accuracy, achieving a rate of 90%. Furthermore, our approach effectively mitigates privacy risks, as evidenced by limiting reidentification accuracies to 47%.

**Conclusions:** This study presents a promising approach to advancing emotion recognition capabilities while addressing privacy concerns in the context of empathetic sensors. By integrating MTL with differential privacy and federated learning, we have demonstrated the potential to achieve high levels of accuracy in emotion recognition while ensuring the protection of user privacy. This research contributes to the ongoing efforts to use affective computing in a privacy-aware and ethical manner.

*JMIR Ment Health* 2024;11:e60003; doi: [10.2196/60003](https://doi.org/10.2196/60003)

**Keywords:** privacy preservation; multitask learning; federated learning; privacy; physiological signals; affective computing; wearable sensors; sensitive data; empathetic sensors; data privacy; digital mental health; wearables; ethics; emotional well-being

## Introduction

### Background

Imagine awakening in the morning to find yourself dealing with stress, perhaps stemming from a disagreement with a

friend, a pressing financial concern, or an unresolved work-related matter. In this scenario, an advanced digital assistant, akin to Siri or Alexa, seamlessly detects the user's stress level through the analysis of physiological signals captured by wearable technology. Leveraging this data, the digital assistant dynamically adjusts its language and tone to suit the

user's mood. Moreover, it proactively recommends personalized relaxation techniques, drawing from past successful interventions such as yoga, mindfulness practices, listening to favorite music, and watching uplifting videos. Such a sophisticated stress recognition and intervention system holds promise for enhancing the well-being and emotional resilience of individuals. Although most of the studies use facial expressions [1] and speech [2] for recognizing stress and emotions (ie, affects as an umbrella term), physiology-based methodologies also emerged as an alternative since they can offer promise for seamless, continuous monitoring within everyday contexts. Notably, 330 million smartwatches, fitness trackers, and wearables were sold in 2020, which are capable of gathering quantitative physiological data [3]. They are promising tools for affect recognition due to their unobtrusive data collection capabilities.

Multitask learning (MTL) is a method proposed to tackle training multiple related tasks at the same time. It works by sharing knowledge between these tasks to make each model perform better [4]. Essentially, MTL acts like a behind-the-scenes helper, enhancing the ability of machine learning (ML) models to generalize different types of data [5]. This technique has been particularly useful in improving ML models across various fields, including affective computing [6]. For instance, if we want to improve how a system recognizes emotions, we can also train it to recognize gender [7]. By doing this, the system can learn from both tasks simultaneously, making it more effective in recognizing emotions and enhancing performance. The motivation for using identity and emotion has also a psychological basis. Connolly et al [8] examined whether facial responses are shared between identity and emotion. They showed that there is a strong positive correlation between face emotion recognition and face identity recognition tasks, and the 2 recognition tasks share a common processing mechanism. Sheng and Li [9] also confirmed the dependence of identity and emotion tasks by using a gait signal. Inspired by the MTL idea and previous studies, we used MTL to perform each task separately, stress recognition and identity recognition, respectively, by using physiological signals. However, there is a concern when third parties gain access to additional information, like gender, identity, or age, which could be misused for targeted advertising or even discrimination in things like job opportunities [10]. Moreover, if sensitive information like biometric data gets leaked, it could lead to cybercrimes such as identity theft. Therefore, it is crucial for MTL systems to incorporate privacy-preserving methods.

Researchers have introduced federated learning (FL) as a way to address privacy issues. FL allows users to train their data locally while sharing only the trained model parameters rather than the original data itself. The philosophy behind FL is to bring code to the data instead of moving the data to the code [11]. This method complies with privacy regulations such as the European Union's General Data Protection Regulation [12] and the California Consumer Privacy Act [13]. FL has shown promise in safeguarding user privacy in Internet of Things networks.

Although FL is a significant step forward in protecting user privacy, it is not invincible. There is still a risk of sensitive information being uncovered by reverse engineering the local model parameters. To further enhance privacy, researchers have integrated privacy-preserving techniques into ML for physiological data. One such method is differential privacy (DP), which adds random noise to each model in the client or server, disrupting updates and limiting the leakage of sensitive information between nodes [14]. However, it is worth noting that DP can reduce the performance of ML models.

In practice, protecting the privacy of users without degradation in model utility is still an open problem. In this study, we first implemented an MTL architecture for recognizing stress and identity from multimodal physiological signals. We further added FL and DP mechanisms to preserve privacy. To obtain a robust performance, we separated the 2 tasks and added noise to only the privacy task, which is biometric identity recognition. In this way, we were able to improve the stress recognition performance with the help of MTL and hide identity information by adding noise to the identity task model by using DP. To the best of our knowledge, this study is the first MTL-based affect recognition study using FL and DP to preserve privacy at the same time.

In the next section, we mentioned the related works using MTL and FL for affect recognition. We then presented our approach in the *Methods* section. In the *Results* section, we compared multitask centralized, decentralized FL, and decentralized FL with DP approaches on selected public datasets for recognizing stress levels and identity of users. We concluded the study with the lessons learned, limitations, and future works in the *Discussion* section.

## Related Works

To develop a robust affective computing system that can be used in practical applications, researchers tested various modalities with state-of-the-art deep learning techniques. MTL has also raised significant attention from various domains over the past few years, including affective computing. Chen et al [14] applied it to audiovisual signals in the Audio/Visual Emotion Challenge and achieved a significant performance increase compared to baseline ML algorithms by detecting arousal and valence levels simultaneously. Since the human face source can be used for several tasks, such as gender recognition or age estimation, Sang et al [7] used MTL with convolutional neural network (CNN) for smile detection, emotion recognition, and gender classification.

They outperformed the state-of-the-art techniques in selected 3 benchmark datasets (Internet Movie Database and Wiki dataset, GENKI-4K dataset [a dataset containing images labeled with facial expressions, commonly used in emotion recognition research], and Facial Expression Recognition Challenge-2013 dataset). MTL-based centralized learning (CL) for affective computing is beneficial for simultaneously sharing features and learning auxiliary tasks. However, when private tasks (ie, face, gender, and person detection) are included in MTL to improve affect recognition performances,

the models create the risk to reveal this sensitive information to possibly malicious parties especially while uploading the features to the central server for both utility and auxiliary tasks for learning.

Data privacy has become an issue of great concern in affect recognition using either verbal or nonverbal data, as gender, age, and identity of the user can be revealed in the process. FL is proposed to preserve privacy while taking advantage of ML and has attracted significant attention from various domains over the past few years, but affective computing research and applications on emotion-related tasks are rarely discussed. Most existing works are conducted on private datasets or in limited scenarios, making it difficult for researchers to compare their methods fairly [15].

FL has been widely applied for facial features for affect recognition. Somandepalli et al [16] investigated FL for 2 affective computing tasks: classifying self-reports and perception ratings in audio, video, and text datasets. Using the speech modality, Latif et al [17] investigated an FL approach in emotion recognition tasks while sharing only the model among clients. They implemented a long short-term memory classifier and tested it on the interactive emotional dyadic motion capture dataset with 4 emotions: happy, sad, angry, and neutral. Face and speech modalities were also often combined to get a more robust performance [18]. Chhikara et al [19] combined emotion recognition from facial expressions and speech with an FL approach. For the face modality, they

used a combination of CNN and support vector machine models, whereas for the audio modality, they applied a 2D CNN model to extract spectrogram images. Their proposed framework has been validated and tested on 2 datasets, Facial Expression Recognition 2013 for facial emotion recognition and Ryerson Audio-Visual Database of Emotional Speech and Song for speech emotion recognition, respectively.

On the contrary, FL has been seldom used in the context of affect recognition from physiological signals, as indicated in Table 1. Can and Ersoy [15] implemented an FL learning model to forecast perceived stress using physiological data. Each subclient uses a multilayer perceptron classifier to locally train its data on the edge, and the sharing of individual updating parameters of multilayer perceptron is facilitated using the federated averaging (FedAvg) algorithm. FL has also been extended to handle multimodal physiological signals. Nandi and Xhafa [20] proposed Federated Representation Learning for Multi-Modal Emotion Classification System, an FL framework-based ML model for emotion recognition. They applied wavelet feature extraction and a neural network to electrodermal activity (EDA) and respiration data from the Dataset for Emotion Analysis using Physiological signals to recognize valence arousal levels, validating their approach. These studies demonstrate the successful application of FL without compromising affect recognition performance.

**Table 1.** Studies using MTL<sup>a</sup> or privacy-preserving approaches for various applications.

Study	Application	Signal	MTL	FL <sup>b</sup>	DP <sup>c</sup>	Multimodality
Zhao et al [21]	Network anomaly detection	Network signals	✓	✓		
Smith et al [22]	Activity recognition	Acceleration, gyroscope	✓	✓		✓
Somandepalli et al [16]	Affect recognition	Speech, video, text		✓		✓
Sang et al [7]	Affect recognition	Face images	✓			
Shome and Kar [23]	Affect recognition	Face images		✓		
Chhikara et al [19]	Affect recognition	Face and speech		✓		✓
Feng et al [24]	Affect recognition	Speech		✓	✓	
Can and Ersoy [15]	Affect recognition	PPG <sup>d</sup>		✓		
Nandi and Xhafa [20]	Affect recognition	EEG <sup>e</sup> , EDA <sup>f</sup>		✓		✓
Wang et al [18]	Biometric	Face speaker		✓	✓	✓
This study	Affect recognition	PPG, EDA, ACC <sup>g</sup> , ST <sup>h</sup>	✓	✓	✓	✓

<sup>a</sup>MTL: multitask learning.

<sup>b</sup>FL: federated learning.

<sup>c</sup>DP: differential privacy.

<sup>d</sup>PPG: photoplethysmography.

<sup>e</sup>EEG: electroencephalography.

<sup>f</sup>EDA: electrodermal activity.

<sup>g</sup>ACC: acceleration.

<sup>h</sup>ST: skin temperature.

While FL has been introduced to enhance model training in terms of privacy, the privacy vulnerabilities inherent in the stochastic gradient descent (SGD) algorithm remain unresolved. DP mechanisms have primarily been discussed in FL settings, involving the injection of noise into each model client or server to perturb updates and limit gradient

leakage shared among nodes (ie, client and server) [25]. In one of the initial applications, authors introduced a new private training method termed differential private SGD, which reduces local and global gradient information leakage between the client and server. Instead of using the standard composition theorem to calculate the final distribution of

overall noise clients, they used a moments accountant metric to adaptively monitor the overall privacy loss. Recognizing that servers are often curious or untrustworthy, Wei et al [26] proposed a local DP mechanism algorithm by introducing Gaussian noise distribution into user models before uploading them to servers. To address the communication overhead required for optimal convergence upper bound for DP, they introduced a novel approach known as communication rounds discounting method, which achieves a better trade-off between the computational complexity of searching and convergence performance. DP has also been leveraged for affect recognition. Feng et al [24] used user-level DP to mitigate privacy leaks in FL for speech emotion recognition. Recently, Smith et al [22] showcased promising performance in preserving privacy via multitask federated learning (MFL) for activity recognition.

When examining the affective computing literature, we find studies using MTL to improve performance and others

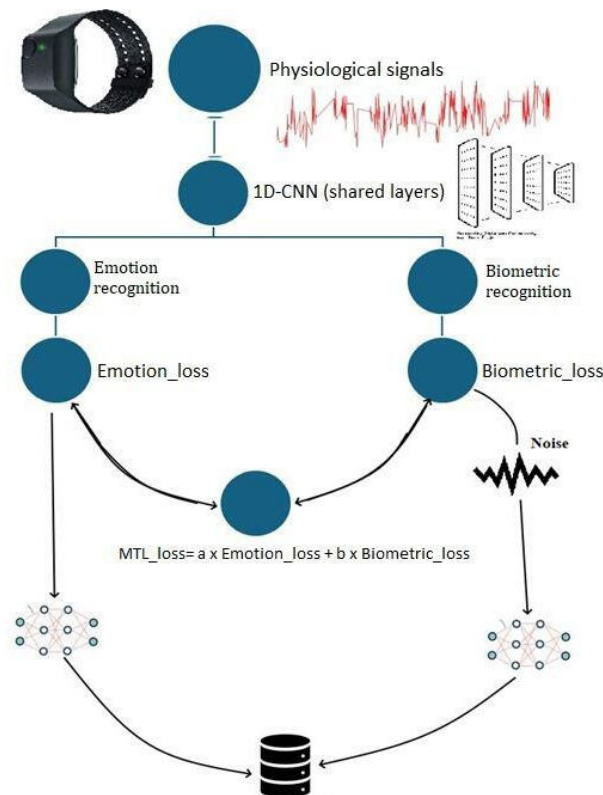
using FL or DP to preserve privacy. However, there are no studies that leverage MTL for performance enhancement while also preserving privacy in this literature. The overall idea of our approach is to separate the utility and the privacy tasks preventing DP from compromising the performance of the utility task (affect recognition) and introducing noise exclusively to the privacy task (biometric identity recognition). Consequently, our work represents the first MTL-based approach to affect recognition using FL and DP to preserve privacy at the same time.

## Methods

### Overview

Our proposed framework is divided into 3 main substeps: feature extraction, FL model, and FL with DP settings (see Figure 1 for the block diagram).

**Figure 1.** Block diagram of the proposed MTL system. CNN: convolutional neural network; MTL: multitask learning.



### Ethical Considerations

This study did not require formal ethics approval, as it exclusively used publicly available datasets, which were obtained from open-access repositories with no personally identifiable information. According to the General Data Protection Regulation, studies involving the use of anonymized, publicly accessible data that do not engage human participants directly are exempt from ethics review. Therefore, no ethics approval was sought or required.

### Data Description and Feature Extraction

For our experiments, we use 2 datasets, wearable stress and affect detection (WESAD) [27] and virtual environment for real-time biometric interaction and observation (VERBIO) [28], which have been created for affective state monitoring.

#### WESAD Dataset

In the WESAD dataset, each participant recorded physiological signals such as blood volume pulse, electrocardiogram (ECG), EDA, electromyogram, respiration, body temperature, and 3-axis acceleration measured from the chest and wrist

using Plux RespiBAN and Empatica E4 devices. In total, 15 (12 male and 3 female) people participated in this experience. There were 4 states: baseline, amusement, stress, and meditation. More details can be found in Schmidt et al [27].

## VERBIO Dataset

The VERBIO dataset [28] consists of biobehavioral responses and self-reported data during public speaking presentations, both in front of real-life and virtual audiences. In total, 55 participants conducted 10 distinct presentations, each lasting approximately 5 minutes, across the 3 segments of the study, spanning 4 days: PRE (1 session, day 1), TEST (8 sessions, days 2-3), and POST (1 session, day 4). The PRE and POST segments involved real-life audiences, while the TEST segment featured various virtual audiences. To prevent participant exhaustion, the TEST segment was divided into 2 days, each comprising 4 sessions. In total, the study yielded 10,800 minutes of acoustic and physiological data from 82 real and 216 virtual reality presentations. In total, 35,174 segments were created. We solely used physiological data collected using the Empatica E4 device, possessing identical sampling rates and modalities as those used in the WESAD dataset.

## Preprocessing

Each modality of the signal is segmented using 700 sample windows size with 50% overlap, as suggested in the literature [29]. To maximize the correlation among interparticipants and minimize among participants, these segments were further processed for extracting features [15] such as mean, variance, root mean square, frequency domain features, average first amplitude difference, second amplitude difference, skewness, kurtosis, and entropy as a nonlinear feature.

## Decentralized MFL Model

For privatizing the user identity while preserving stress recognition accuracy, we adopted an MFL approach that can effectively improve the performance of stress recognition while limiting the risk of inferring sensitive information from the training model since the client does not want to be exposed to the cloud service provider. MFL architecture-based stress recognition is developed as follows:

1. The dataset is partitioned into  $K$  clients. The data size of all the clients is the same. The client distribution is also assumed as independent and identically distributed (IID) and no IID [10].10
2. For the local training process, there is only 1 iteration for SGD local training for each client. In particular,  $\mathbf{w}$  is the local model parameter [10] given by:

$$\mathbf{w}_U^{D_i} = \operatorname{argmin}_{\mathbf{w}_U} \left( F_U(\mathbf{w}_U) + \frac{\mu}{2} \left\| \mathbf{w}_U - \mathbf{w}^{(D_i-1)} \right\|^2 \right) \quad (1)$$

where  $\mu$  is the learning rate and  $F_U$  is the local loss function of the  $k$ th user.

3. The local data of different clients cannot be communicated, and only the models can be shared.1010

4. Following the FedAvg algorithm [10], the server uses a global averaging approach to aggregate all local training models to compute the final global model. Formally [10], the server aggregates the weights sent from the  $K$  clients as follows:

$$\mathbf{w} = \sum_{U=1}^K p_i \mathbf{w}_U^{D_i} \quad (2)$$

where  $w_i$  is the parameter vector trained at the  $k$ th client,  $w$  is the parameter vector after aggregating at the server,  $K$  is the number of participating clients,  $D_i$  is the dataset size of each participating client,  $D = UD_i$  is the whole distributed dataset, and  $P_i = |D_i|/|D|$ .

5. The global training epoch is set to  $M$  rounds (aggregations). The server solves the optimization problem [10]:

$$\mathbf{W}^* = \operatorname{argmin}_{\mathbf{w}_U} \sum_{U=1}^M P_U F_U(\mathbf{w}_U, \mathcal{D}_i) \quad (3)$$

where  $F_U$  is the local loss function of the  $k$ th client. Generally, the local loss function is given by local empirical risks.

## Decentralized FL With DP

In conventional FL, the global model is computed through averaging over model client participants, which performs better within homogeneous FL settings. However, using inference or adversarial attack, this shared model may contain sensitive and private information such as gender, age, and biometric template user. In such cases, the MFL framework is required to reduce the leakage of the black box gradient exchanged model. To overcome this limitation, researchers have used the DP scheme to protect either local or global data training FL model. However, the perturbed gradient using DP with a low budget has high variance, which leads to worse performance and slower convergence. We also compared adding noise to whole model and task-specific last layers in addition to the shared layers and demonstrated the performance of these 3 different approaches (Multimedia Appendix 1). Motivated by personalized FL [26], our work focuses on client-level privacy, which aims at a private specific layer of the client model rather than perturbing the entire whole local model. This is because the base layers are mostly redundant information, while the most important information that holds private and public information is located in the upper layers. To meet the utility privacy trade-off, the DP mechanism is used to perturb the gradients using Gaussian noise at a specific layer or task. Here, we use all steps in the FL model except step 4, that is, before uploading the local SGD model client to the global server, we inject an amount of noise to the updated local parameters. In that sense, we perturb the local gradient training inference with two kinds of noise distributions:

1. An additive Gaussian noise  $\eta \sim N(0, \sigma)$  to each weight. This operation can be mathematically described as follows:

$$w_{t+1} = w_t + \eta \quad (4)$$

2. A set of noise distributions can be sampled from the DP mechanism. A randomized mechanism  $\mathcal{M}$  on the training set with domain  $X$  and range  $R$  satisfies  $(\epsilon, \delta)$  – DP for 2 small positive numbers and if the following inequality holds [25]:

$$Pr[\mathcal{M}(x) \in \mathcal{S}] \leq e^\epsilon Pr[\mathcal{M}(x') \in \mathcal{S}] + \delta \quad (5)$$

where  $x$  and  $x' \in X$  are 2 input neighbor datasets, and  $\mathcal{S} \subseteq R$ , where  $R$  is the set of all possible outputs,  $\delta$  is privacy loss or failure probability, and  $\epsilon$  is privacy budget.

An ideal DP mechanism provides a lower value of  $\delta$  and a smaller value of  $\epsilon$ . Unfortunately, these values decrease the function utility (eg, accuracy metric), so the main question is how much DP values we must perturb its output while guaranteeing trade-off privacy-utility. Intuitively, an output perturbation mechanism takes an input  $x$  and returns a random variable  $s(x)$ . This operation can be modeled by:

$$\mathfrak{M}(x) = s(x) + N_\sigma \quad (6)$$

where  $N$  is scaling noise sampled from a specific distribution. In this work, we chose Laplace and Gaussian mechanisms [25] that use L1 and L2 norm sensitivity, respectively. The sensitivity function can be expressed as:

$$\Delta f = \max_{D, D'} \|s(x) - s(x')\|_{1,2} \quad (7)$$

Scaling noise can be computed as:

$$\sigma = \Delta f / \epsilon \quad (8)$$

Output perturbation satisfies  $(\epsilon, \delta)$  – DP when we properly select the value scaling noise. Thus, it sampled from Laplace and Gaussian distributions [25] as:

$$M_{\text{Laplace}}(x, f, \epsilon, \delta) = s(x) + \text{Lap}(\mu = 0, b) \quad (9)$$

$$M_{\text{Gaussian}}(x, f, \epsilon, \delta) = f(x) + \mathcal{N}(\mu = 0, \sigma^2) \quad (10)$$

The gradient information leakage can be reduced by applying a gradient thresholding or clipping algorithm. As explained by Abadi et al [11], gradient clipping is crucial in ensuring the DP of FL algorithms. So, each provider or client model update needs to have a bounded norm, which is ensured by applying an operation that shrinks individual model updates when their norm exceeds a given threshold. Clipping impacts of an FL algorithm's convergence performance should be known to create FL algorithms that protect DP.

## Results

### Overview

Three scenarios are created to tackle the aforementioned challenges with the DP learning approaches: centralized, decentralized FL, and decentralized FL with DP. Their performances are evaluated on WESAD and VERBIO datasets on 2 different tasks. The first task is identifying users from a set of registered and recorded users. The second task is perceived binary stress recognition, which tries to distinguish the user's stress level, stress versus nonstress. We train a multitask deep learning model for handling these tasks simultaneously. In addition, for better training and to avoid overfitting, an early stopping regularization technique is used as gradient descent. The accuracy metric is used for measuring identification and stress recognition performance. In each simulation scenario, we run 5-fold cross-validation, where each fold is tested based on the training of the other 4. Inspired from a successful architecture [30] and as described in Table 2, the multitask 1D-CNN model is based on 3 convolutional layers, a pooling layer, 2 fully connected layers, and 2 linear classifiers to classify the studied tasks. The multitask model uses the cross-entropy loss function and SGD learning rate ( $\beta=.0005$ ).

**Table 2.** Hyperparameters of the multitask 1D-convolutional neural network–based architecture.

Layer	Type	Hyperparameters
Input	Input	Features size
Conv1D	Convolution	Input=1, output=20, K=8, stride=1, padding
ReLu	Activation function	ReLu
Pooling	Pooling	Stride=2, max pooling
Conv	Convolution layer	Input=20, output=40, K=8, stride=1, padding
Relu	Activation function	ReLu
Pooling	Pooling	Stride=2, max pooling
Conv	Convolution layer	Input=40, output=60, K=8, stride=1, padding
Relu	Activation function	ReLu
Pooling	Pooling	Stride=2, max pooling
Fully connected 1	Fully connected layer	Input=360, output=100
Fully connected 2	Fully connected layer	Input=100, output=300
Linear 1	Input: 100, output: 2	Output=2 affect classes, activation function:linear
Linear 2	Input: 300, output: N	Output=N identity classes, activation function:linear

For each target task, the individual loss is determined by the cross-entropy for both stress recognition ( $Loss_1$ ) and identification tasks ( $Loss_2$ ). The individual losses are summed and form the total cost function ( $Loss_T$ ).

## CL Approach

Most existing stress recognition studies in conventional centralized ML settings use WESAD to position their works against state of the art [31-33] and achieve over 95% accuracy for binary stress recognition. Similarly, we used the CL approach on the WESAD dataset as a baseline experiment. We measured the performance of stress levels for each participant independently and then computed the average performance accuracy. Each individual participant model is trained using 5-fold cross-validation.

Multimedia Appendix 2 shows the results of our CL approach using a 1D-CNN multitask model for stress and identity recognition tasks on 2 datasets: WESAD and VERBIO. After training, the output layers are used to infer the stress level and the identity of the user, resulting in quite similar average scores of 99%. The results reveal a potential information leakage in this case, wherein model accuracy is preserved at the expense of the user's privacy. As a result, this approach fails to ensure the users' privacy since their data are transmitted to the server for training purposes.

## Multitask Decentralized FL Approach (MFL)

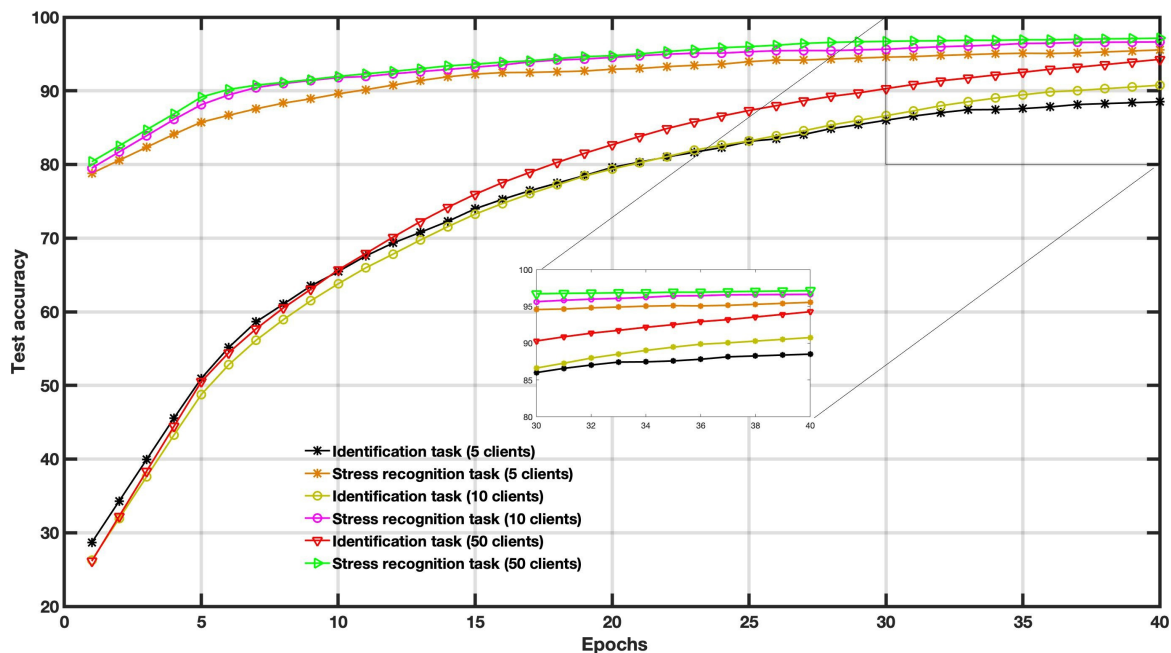
Here,  $K$  (ie, participating clients) training models were created to train the whole dataset and the size of local samples

$D_i=1000$ . We set the number of training epochs (communication round) to  $T=40$  and local training epochs to 1. Overall, the average accuracy result achieved is 97% and 95% for stress mood recognition and 93% and 90% for user identification on WESAD and VERBIO, respectively.

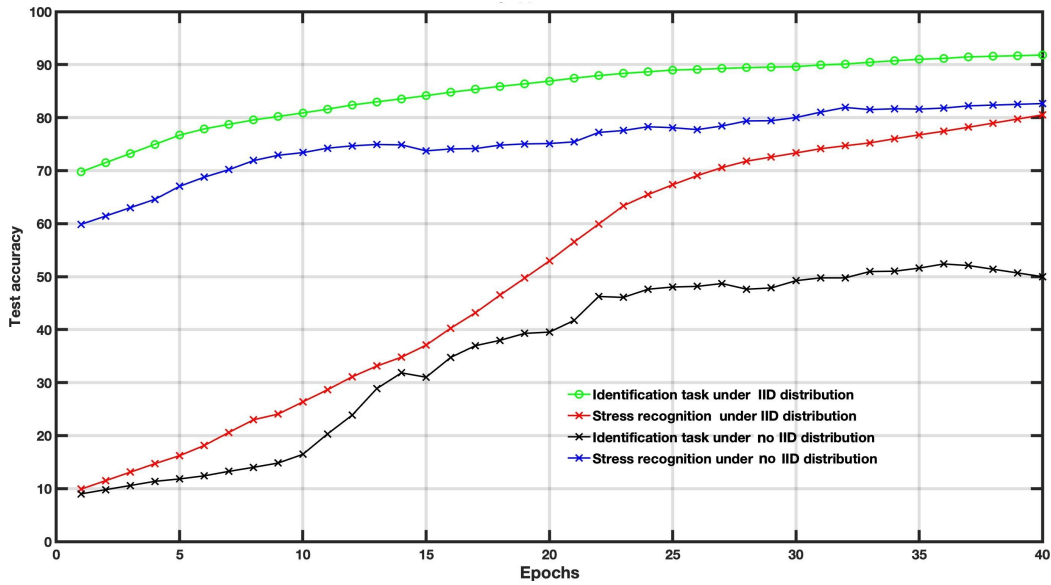
To examine the effect of client participation within the MFL model, we tested different numbers of clients, that is,  $K=5$ ,  $K=10$ , and  $K=20$ . As reported by Wang et al [34] and confirmed in Figure 2, an increasing number of clients and more client participation provide better performance for MFL training. The client distribution is different in assessing the MFL model in real-world conditions. We compare the convergence performance of the MFL model under IID and no IID for both tasks: stress recognition and subject identification (Figure 3). We note that the data distribution dramatically affects the quality of the FL training and obviously affects MFL's convergence performance. IID is a more idealistic case, whereas no IID is a more realistic case. Therefore, we expect the result to be lower in no IID cases. The performance difference proves the importance of data distribution.

Adjusting FL hyperparameter setting results can achieve a better performance than the CL approach. However, it may lead to a lower privacy level. As a result, SGD training may still reveal sensitive information about the client while exchanging the ML model with the global server.

**Figure 2.** The impact of the user participant size (various number of clients) on the multitask federated learning performance (wearable stress and affect detection).



**Figure 3.** The impact of IID versus no IID distribution on the multitask federated learning performance. IID: independent and identically distributed.



### MFL With DP Approach

- To highlight the benefits of our proposed approach, we examine the impact of injecting noise into the local client training network according to these 3 scenarios: the full layers, shared layers, and task-specific layers (Figure 2). The used noise is sampled via the following mechanisms:
- With DP technique-based Laplace and Gaussian mechanisms, the noise scale is drawn from the output perturbation mechanism. DP parameters are computed at each local training round to generate appropriate noise injected from specific distributions (ie, Laplace and Gaussian). Besides appropriate  $(\epsilon, \delta)$ -DP initialization, there are a few hyperparameters to be tuned, such as the number of clients  $N$ , the number of maximum communication rounds  $T$ , and the number of chosen clients  $K$ .
  - Laplace distribution [11] is computed as:

$$\sigma = \frac{\Delta f}{\epsilon} \tag{11}$$

- Gaussian distribution—2 distributions are given by Abadi et al [11] and Schmidt et al [27], respectively:

$$\sigma_1 = \sqrt{\frac{2 \log \frac{1.25}{\delta}}{\epsilon}} \tag{12}$$

$$\sigma_2 = \frac{\Delta f \sqrt{2qT \log(\frac{1}{\delta})}}{\epsilon} \tag{13}$$

where  $\Delta f = \frac{2C}{U_i}$ ,  $q = \frac{K}{N}$ , and  $C =$  clipping threshold. We set the clipping factor to 1 and  $\delta$  to 0.00001.

Figures 4-6 show the accuracy comparison when adding Gaussian distribution levels into local training according to the 3 scenarios on the WESAD dataset. Compared to the baseline scenario, that is, no private mechanism, the perturbing share layer scheme with  $\sigma=0.1$  and  $\sigma=0.3$  only provides better results for utility tasks; however, the identification task reached an accuracy of around 86%.



Figure 4. Decentralized multitask federated learning approach results under different Gaussian noise levels ( $\sigma=0.1$ ).

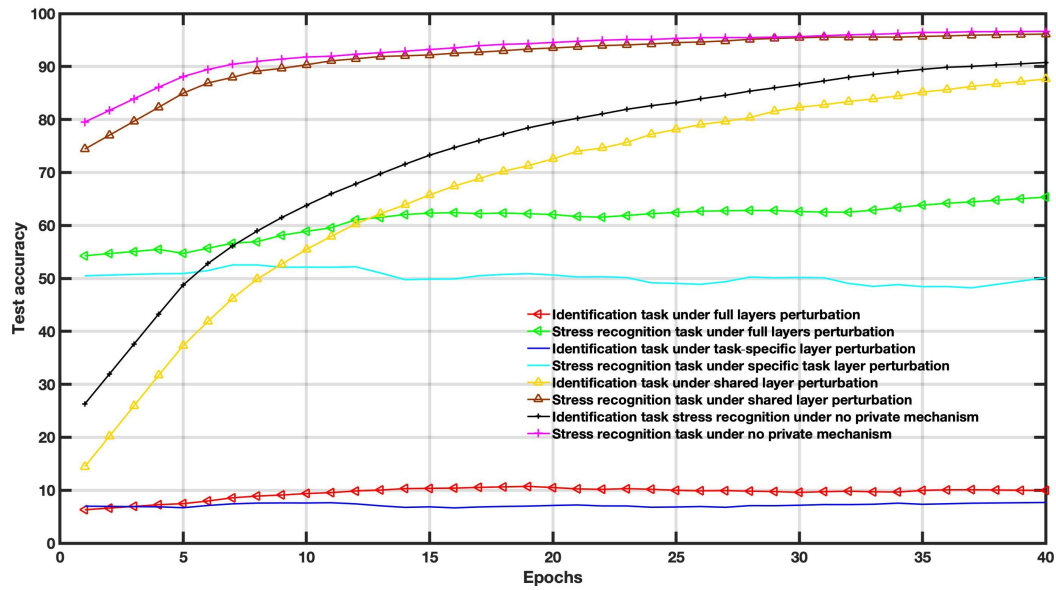
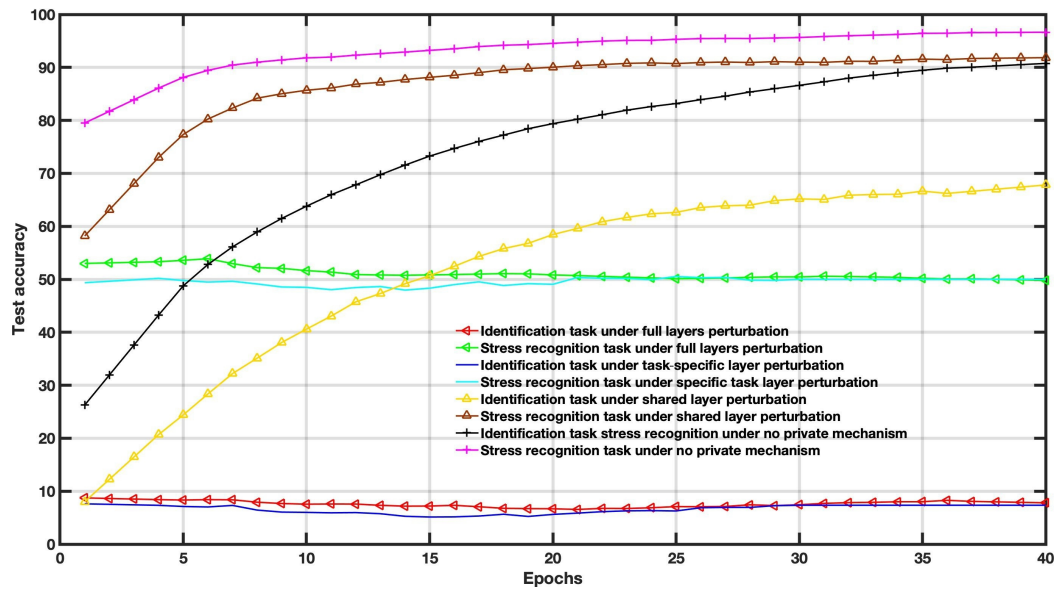
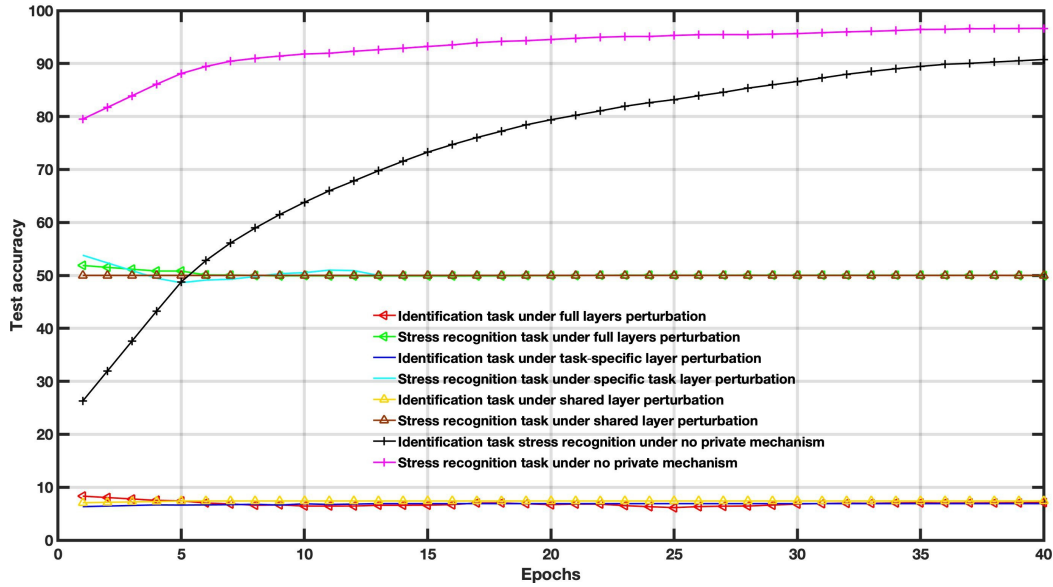


Figure 5. Decentralized multitask federated learning approach results under different Gaussian noise levels ( $\sigma=0.3$ ).



**Figure 6.** Decentralized multitask federated learning approach results under different Gaussian noise levels ( $\sigma=0.6$ ).

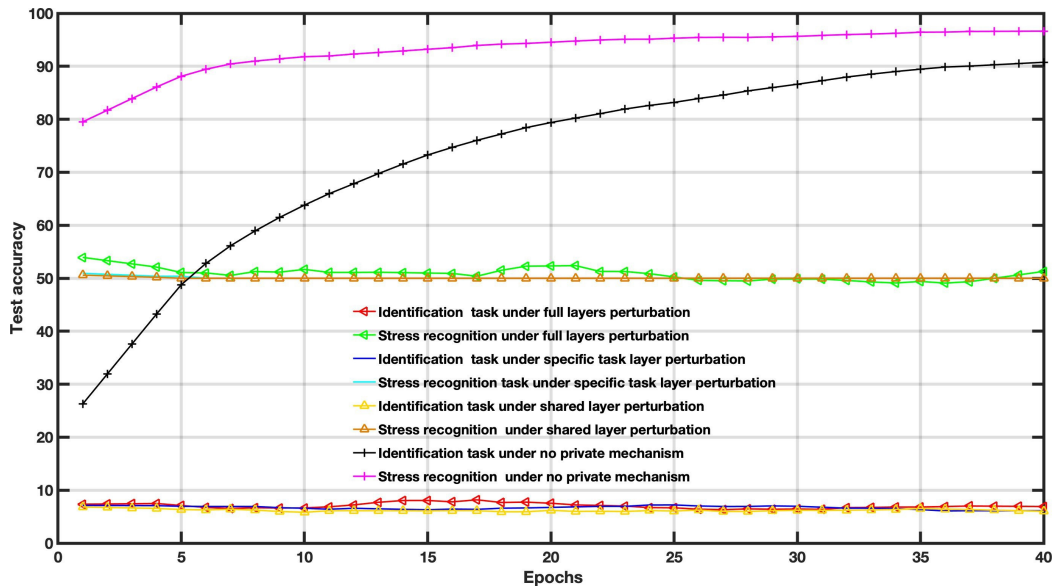


In this case, the amount of noise drawn from the Gaussian distribution is used to balance the utility and privacy and does not consider the FL settings parameters. The Gaussian levels are set to  $\sigma=0.1$ ,  $\sigma=0.3$ , and  $\sigma=0.6$ . For instance, increasing Gaussian noise leads to poor performance, as depicted in Figure 6.

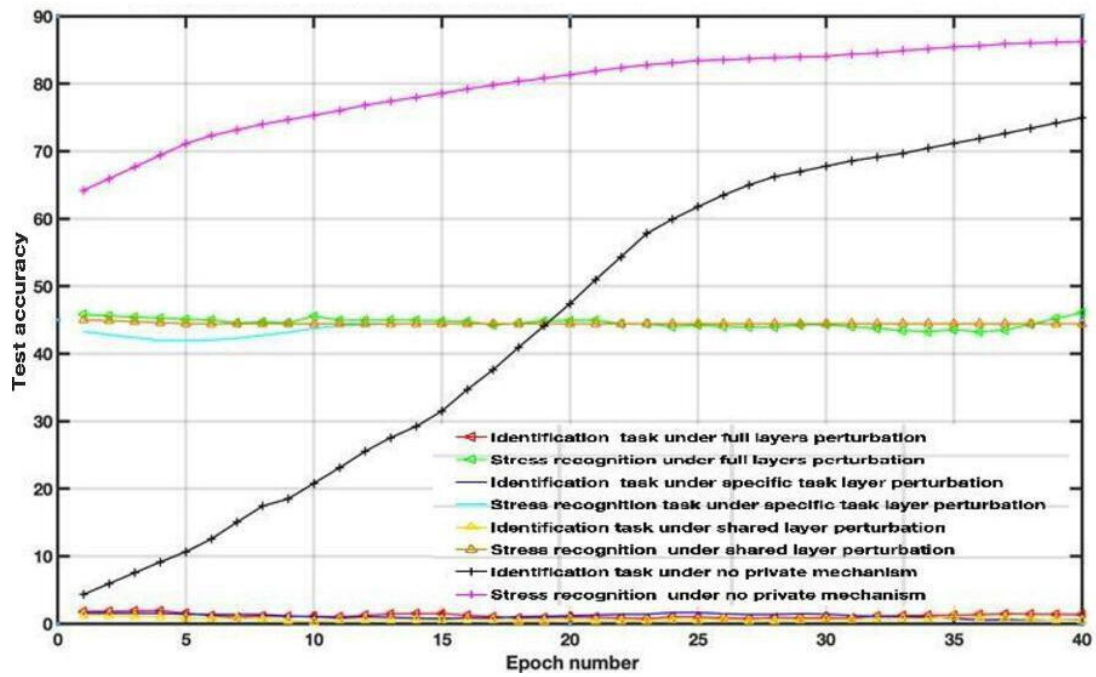
To examine the DP impact on the utility-privacy trade-off, we assessed the performance of MFL with a DP mechanism in WESAD and VERBIO datasets under the aforementioned scenarios. The DP budgets are set as follows  $\epsilon=1$ ,  $\epsilon=15$ ,

$\epsilon=50$  and  $\epsilon=2000$  for this experiment. As depicted in Figures 7-14 compared to the baseline scenario, that is, no privacy enhancing mechanism, adding DP noise into both shared and specific layers provides better results for utility performance; however, in terms of privacy, the perturbing specific task layer scheme provides better results than the perturbing shared layer. Results show that FL with perturbing all layers slows the convergence compared to others, although it provides better privacy (ie, decreasing identification accuracy).

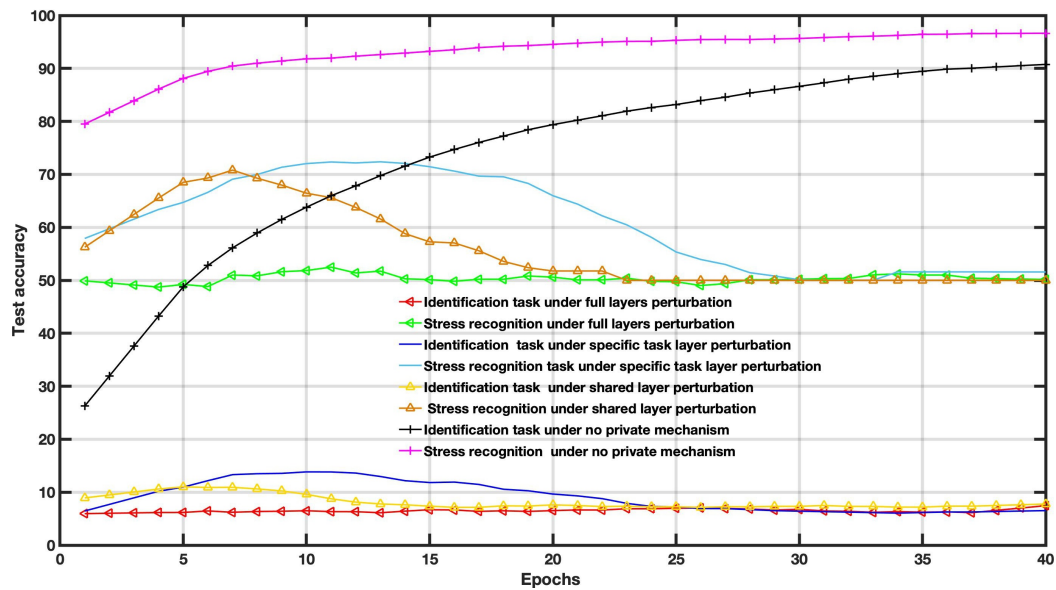
**Figure 7.** Multitask federated learning performance under the differential privacy mechanism “different layer perturbations” (wearable stress and affect detection,  $\epsilon=1$ ).



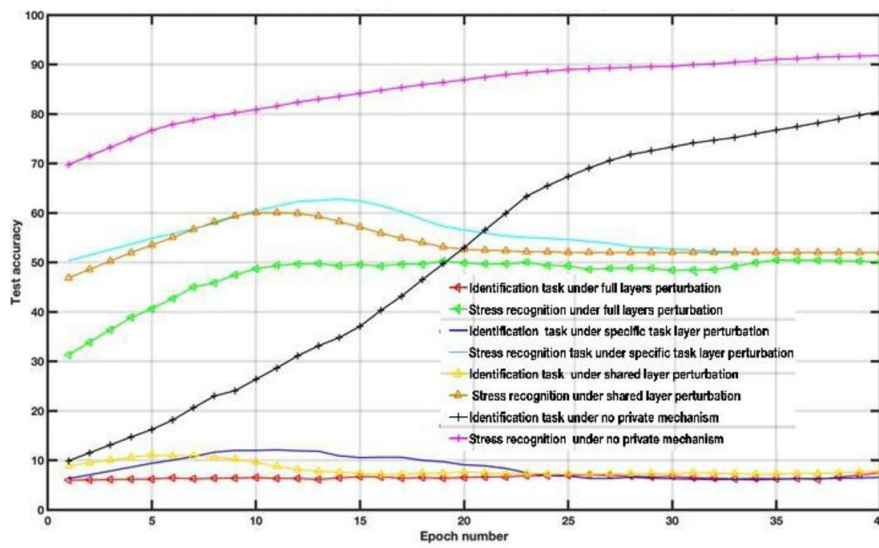
**Figure 8.** Multitask federated learning performance under the differential privacy mechanism “different layer perturbations” (virtual environment for real-time biometric interaction and observation,  $\epsilon=1$ ).



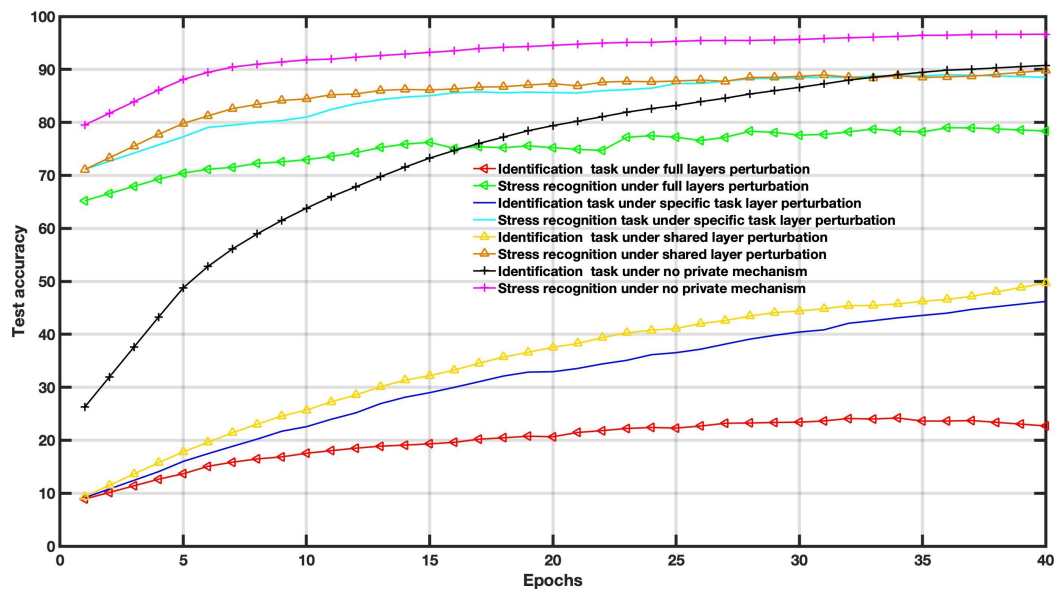
**Figure 9.** Multitask federated learning performance under the differential privacy mechanism “different layer perturbations” (wearable stress and affect detection,  $\epsilon=5$ ).



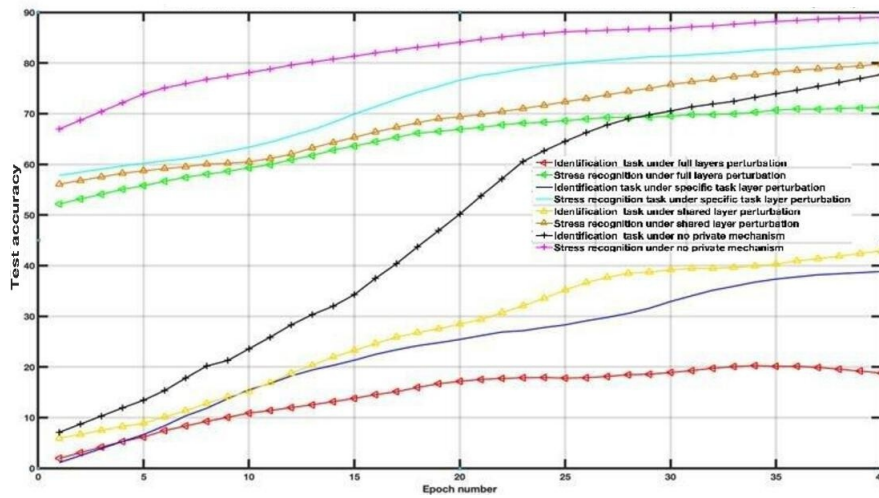
**Figure 10.** Multitask federated learning performance under the differential privacy mechanism “different layer perturbations” (virtual environment for real-time biometric interaction and observation,  $\epsilon=5$ ).



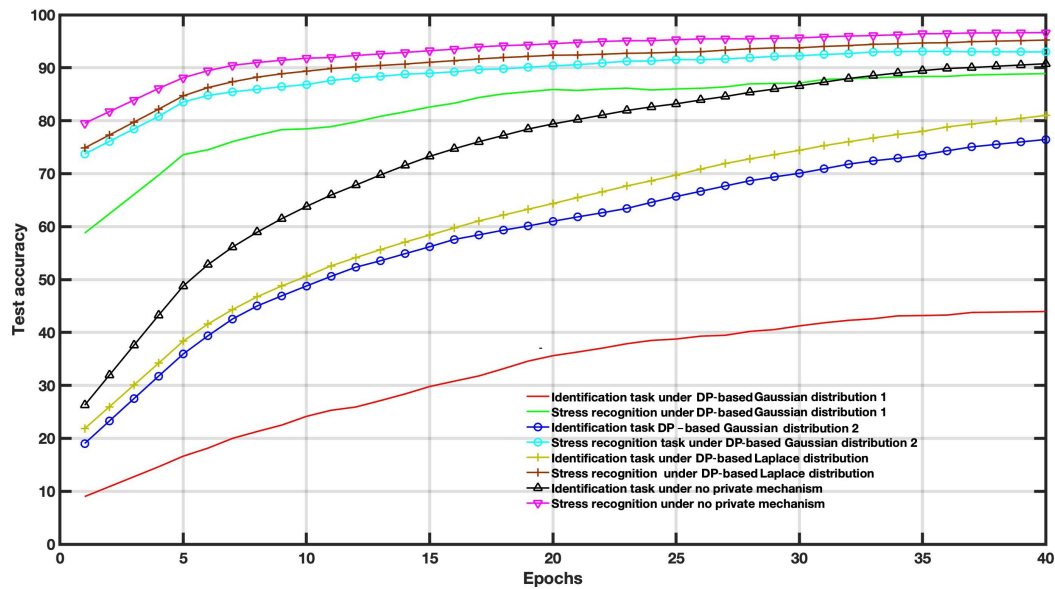
**Figure 11.** Multitask federated learning performance under the differential privacy mechanism “different layer perturbations” (wearable stress and affect detection,  $\epsilon=15$ ).



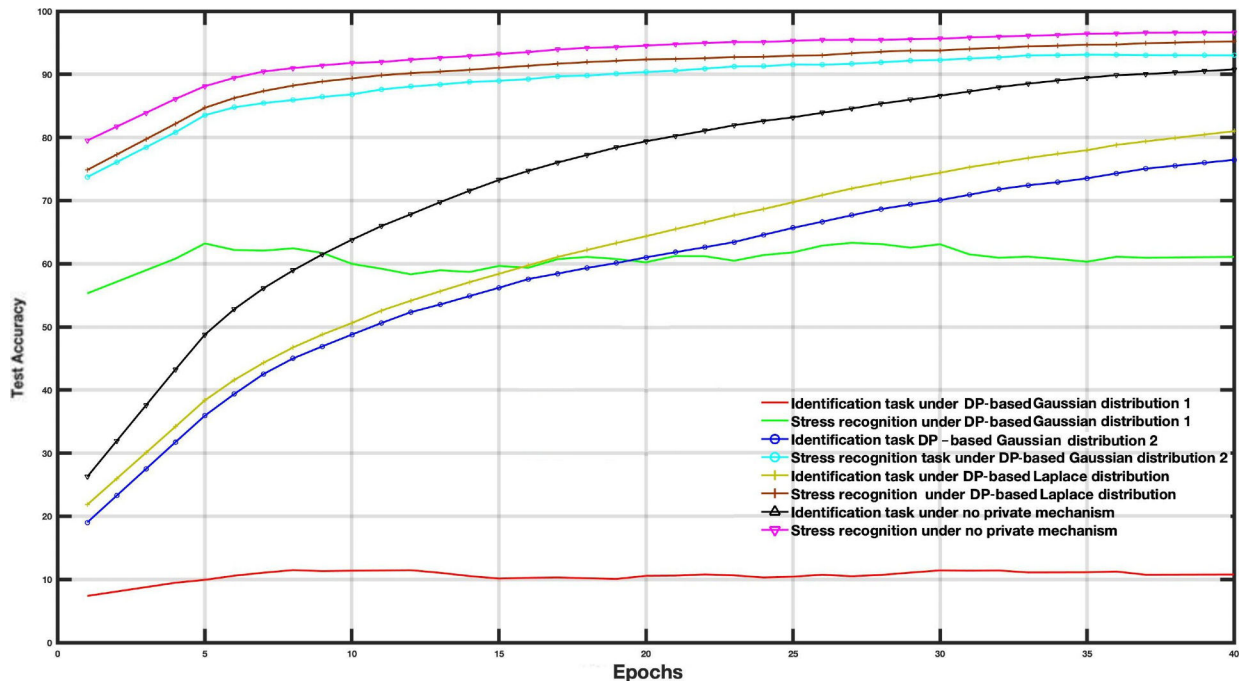
**Figure 12.** Multitask federated learning performance under the differential privacy mechanism “different layer perturbations” (virtual environment for real-time biometric interaction and observation,  $\epsilon=15$ ).



**Figure 13.** The evaluation results of the multitask federated learning approach with 3 differential privacy (DP) distributions (various layer task perturbation,  $\epsilon=15$ ).



**Figure 14.** The evaluation results of the multitask federated learning approach with 3 differential privacy (DP) distributions (full perturbation,  $\epsilon=15$ ).



Intuitively, our results demonstrate that adding noise to upper layers (identity recognition layers) effectively achieves a better privacy-utility trade-off. This advantage comes at the expense of a formal quantification of the relationship between learning features, that is, what we aim to share, and private variables, that is, what we aim to protect, which is rarely available in practice. We also evaluate the impact of noise distribution type on the proposed framework performance. The results from the WESAD dataset demonstrate that adding Laplace noise in our local training model preserves the stress recognition accuracy better than the Gaussian noise types (Figures 13 and 14). However, it also maintains the identity recognition task performance.

Nevertheless, using the Gaussian mechanism (ie, Gaussian 2) increases the privacy level of the local training model because increasing the number of global iterations will also negatively affect its global convergence performance, that is, a larger  $T$  would increase the noise level variance, dramatically decreasing the global accuracy model (equation 13). In addition, we have found that a larger  $K$  contributes to avoiding the vanishing local SGD gradient problem; however, a larger  $N$  leads to a scale-down in the variance of injected noise level to model parameters and fools the SGD training inference.

When DP is also used in FL settings, our experiments suggest that it achieves encouraging performance even on lower budget values (ie, increasing privacy requirements). The compromise of physiological data privacy can have significant consequences for a life of individuals. It provides an opening for data attackers and puts them at danger of data breaches, which can result in several threats [35]. For instance, Shen et al [36] provided an analytic study to estimate the leakage level of the global training FL model about user's information, and they generated a membership attack system on the global model to check if the target user

is participated in training FL model or not. The experiences have been validated and tested on health data, called the University of Michigan Intern Health Study. The authors suggested to use DP to protect user-level privacy while maintaining the utility task performance. In similar work [37], applied personalized FL in recognizing stress using mobile health data (ie, WESAD) to decrease the information leakage from the global FL model.

Besides, there are also other risks including the possibility of revealing a soft biometric (eg, gender, location, and authenticity) and a hard biometric (ie, identity). So far, most of the research in using physiological modalities in both biometric domains focused on ECG, electroencephalography, EDA, and photoplethysmography [35,38]. We found that the model obviously achieves a better balance between the stress and identification classification task when multimodal signals are combined and used as an input compared to the single modality. However, as aimed in our work, it is difficult to provide privacy guarantees since there is no study that provides an investigation on which modality can reveal sensitive information about the subject. In this experiment, we provide a comparative study among these used modalities for both tasks, subject identification and emotion recognition. According to Tables 3 and 4, we found that the ECG and acceleration modalities always led to an increase in the performance of the identification task. This fact might have contributed to their popularity in the biometric field [38]. This observation hints us that discarding model parameters containing sensitive information learned from modalities associated with identity information might help to effectively achieve a good privacy-utility trade-off to a certain extent, that is, decrease information leakage. Although for these modalities (eg, ACC and ECG for WESAD), the identification accuracy is higher than binary stress recognition accuracy, in general, it is lower than stress recognition

performance. We can interpret this fact by the number of labels and the complexity of the task. For WESAD, there are 15 participants and for VERBIO, there are 55 participants, which corresponds to 15 and 55 different labels for biometric identification ML systems. If we think that the signals of some people can be similar under the same experimental

conditions, it can cause some confusion for the ML algorithm and can cause some performance decrease. On the other hand, for stress, we used binary labels, and although the stress response can change from individual to individual, it can be stated that it is a relatively simpler task with 2 labels.

**Table 3.** The effect of modality on the stress and subject identification performance (wearable stress and affect detection).

Modalities and task	Approach								
	Centralized			FL <sup>a</sup>			FL+DP <sup>b</sup>		
	ACC <sup>c</sup>	F <sub>1</sub>	AUC <sup>d</sup>	ACC	F <sub>1</sub>	AUC	ACC	F <sub>1</sub>	AUC
<b>ACC</b>									
Stress	91.60	86.09	92.01	89.90	84.81	90.35	84.46	79.23	85.00
Subject identification	95.50	91.70	96.00	94.88	90.76	96.34	58.19	52.01	60.00
<b>BVP<sup>e</sup></b>									
Stress	98.34	96.14	99.13	96.05	92.01	97.00	90.97	86.92	92.46
Subject identification	94.96	90.88	95.00	92.05	88.03	92.88	54.95	49.89	55.00
<b>EDA<sup>f</sup></b>									
Stress	99.00	98.01	99.00	95.29	91.09	95.57	91.27	86.03	92.00
Subject identification	86.15	70.06	86.88	85.89	80.78	86.27	50.05	43.00	50.05
<b>ECG<sup>g</sup></b>									
Stress	97.01	95.00	97.02	93.37	89.03	94.00	91.60	86.45	92.10
Subject identification	98.62	96.54	99.00	90.96	82.89	91.50	68.20	61.03	69.00
<b>TEMP<sup>h</sup></b>									
Stress	99.00	98.00	99.00	94.88	90.76	95.23	90.00	85.00	90.00
Subject identification	97.88	95.78	98.00	91.22	86.01	92.00	51.23	43.00	52.00
<b>All</b>									
Stress	99.86	96.72	99.86	95.56	91.38	96.13	90.00	85.00	90.00
Subject identification	98.56	96.47	99.00	92.00	88.00	92.00	47.00	40.00	47.00

<sup>a</sup>FL: federated learning.

<sup>b</sup>DP: differential privacy.

<sup>c</sup>ACC: acceleration.

<sup>d</sup>AUC: area under the curve.

<sup>e</sup>BVP: blood volume pulse.

<sup>f</sup>EDA: electrodermal activity.

<sup>g</sup>ECG: electrocardiogram.

<sup>h</sup>TEMP: skin temperature.

**Table 4.** The effect of modality on the stress and subject identification performance (virtual environment for real-time biometric interaction and observation).

Modalities and task	Approach								
	Centralized			FL <sup>a</sup>			FL+DP <sup>b</sup>		
	ACC <sup>c</sup>	F <sub>1</sub>	AUC <sup>d</sup>	ACC	F <sub>1</sub>	AUC	ACC	F <sub>1</sub>	AUC
<b>ACC</b>									
Stress	90.61	82.56	91.50	85.48	80.13	86.45	82.18	77.02	83.03
Subject Identification	86.42	83.00	87.12	84.40	79.16	85.00	63.60	58.46	63.60
<b>BVP<sup>e</sup></b>									
Stress	92.96	88.87	93.00	88.95	83.84	90.56	85.00	80.00	85.00
Subject Identification	85.63	80.31	86.33	80.86	74.37	90.13	48.64	42.51	49.00
<b>EDA<sup>f</sup></b>									
Stress	93.16	89.01	94.56	92.46	88.23	93.00	89.24	84.06	90.01
Subject identification	88.30	83.12	88.30	83.83	78.65	84.23	45.12	38.01	46.00

Modalities and task	Approach								
	Centralized			FL <sup>a</sup>			FL+DP <sup>b</sup>		
	ACC <sup>c</sup>	F <sub>1</sub>	AUC <sup>d</sup>	ACC	F <sub>1</sub>	AUC	ACC	F <sub>1</sub>	AUC
<b>TEMP<sup>g</sup></b>									
Stress	92.64	88.37	93.34	90.42	85.12	90.42	86.32	81.12	87.06
Subject identification	85.34	80.07	86.17	82.56	76.33	82.56	39.59	31.42	40.00
<b>All</b>									
Stress	95.00	91.01	95.00	93.96	89.88	94.00	90.94	85.89	90.94
Subject Identification	90.00	85.01	90.00	88.92	83.84	89.10	38.46	30.18	39.35

<sup>a</sup>FL: federated learning.

<sup>b</sup>DP: differential privacy.

<sup>c</sup>ACC: acceleration.

<sup>d</sup>AUC: area under the curve.

<sup>e</sup>BVP: blood volume pulse.

<sup>f</sup>EDA: electrodermal activity.

<sup>g</sup>TEMP: skin temperature.

## Comparison With State of the Art

The proposed approach is also compared with existing approaches used for stress recognition in FL settings, and the results are reported in Table 5. Most similar works have mainly focused on the benefits of FL in affective computing by generating data clients from unique datasets (ie, as used in the FL framework with standard benchmarks, namely Modified National Institute of Standards and Technology, Canadian Institute for Advanced Research-10, and Canadian Institute for Advanced Research-100). To the best of our knowledge, the number of studies using physiological signals

in FL settings is relatively low compared to other case studies in affective computing, such as speech and facial expression-based studies. According to Table 5, we can see that existing works used only FedAvg with the WESAD dataset and achieved similar performance, as confirmed by our previous study. To the best of our knowledge, our paper both improves privacy protection and provides slightly higher accuracies when compared with similar studies. Consistent with compared works, for both databases, WESAD and VERBIO, the obtained performance behaves similarly.

**Table 5.** Comparison of affect recognition studies using federated learning (FL) and privacy-preserving approaches.<sup>a</sup>

Study	Dataset	FL algorithm	Data split	Accuracy
Almadhor et al [39]	WESAD <sup>b</sup>	FedAvg <sup>c</sup> +logistic regression	N/A <sup>d</sup>	86.82
Fauzi et al [40]	WESAD	FedAvg+DNN <sup>e</sup> network	N/A	85.75
Can and Ersoy [15]	Private dataset	FedAvg+MLP <sup>f</sup>	N/A	88.55
Lee et al [29]	WESAD	FedAvg+MLP	LOOCV <sup>g</sup>	75.00
Our previous study [38]	WESAD	FedAvg+1 DCNN <sup>h</sup> +DP <sup>i</sup>	5-fold CV <sup>j</sup>	90.00
This study	VERBIO <sup>k</sup>	FedAvg+1 DCNN+DP	5-fold CV	88.67

<sup>a</sup>The table reports the main approaches have been applied for stress detection by using a physiological dataset.

<sup>b</sup>WESAD: wearable stress and affect detection.

<sup>c</sup>FedAvg: federated averaging.

<sup>d</sup>N/A: not applicable.

<sup>e</sup>DNN: deep neural network.

<sup>f</sup>MLP: multilayer perceptron.

<sup>g</sup>LOOCV: leave-one-out cross-validation.

<sup>h</sup>DCNN: deep convolutional neural network.

<sup>i</sup>DP: differential privacy.

<sup>j</sup>CV: cross-validation.

<sup>k</sup>VERBIO: virtual environment for real-time biometric interaction and observation.

Unlike the compared works [15,29,38,40], which analyze privacy concerns in stress recognition, we introduced a multitask model enabling us to strike a balance between utility and privacy. This is achieved by selectively adding noise to layers prone to subject information leakage. The comparison provided in Table 5 reveals that prior studies achieved performance comparable to our case study when using FedAvg on WESAD or a private dataset. Through

our experiments, we provided a comprehensive investigation of how to balance privacy preservation and stress recognition. The obtained results demonstrate that the combination of MFL and DP can maintain good performance without sacrificing the user's privacy.



## Discussion

The overall objective of our work was to present an approach to stress recognition that ensures robustness while protecting the user's privacy. To this end, we developed a personalized multitask federated model framework with DP. We used a user-level DP mechanism by injecting an amount of noise into personalized layers for perturbing identity while preserving task-specific utility. Using multitask and DP ( $\epsilon=15$ ), we obtained 90% and 87% accuracy for recognizing emotions while limiting the reidentification accuracies to 47% and 38% on WESAD and VERBIO, respectively. We extensively tested different parameters including layer of neural network, privacy budget, and different noise distributions. As expected, adding noise to the upper layer decreases the affect recognition performance less when compared to the last task-specific layers. We also demonstrated the effect of data distribution on the performance. Having said that, the paper is not without limitations. Although we tested our algorithms with state-of-the-art datasets (WESAD 15 participants and VERBIO 55 participants), especially for the

biometric task, larger datasets with a more heterogeneous and higher number of participants should be used. Furthermore, these datasets were recorded in a controlled environment. Real-life or in-the-wild datasets will create more challenges and might require more advanced and complex architectures. However, with the available data, we managed to develop an automatic stress recognition system with around 90% accuracy by keeping biometric identification accuracy below 40%. We believe that our results will help researchers in determining suitable parameters and distributions for achieving the desired trade-off between utility and privacy for affective computing applications. Currently, new gradient-based unsupervised adversarial attackers are attacking deep neural classification models to infer the privacy of the distributed training gradient. In future works, to address this threat, we are planning to conduct additional experiments with federated differentially private generative adversarial networks that can provide better privacy protection and data diversity for widespread applications of physiological computing systems.

## Acknowledgments

This work was carried out within the framework of the AI Production Network Augsburg. This work was partly funded by the German Research Foundation under the project "Health-relevant effects of different urban forest structures" (Learning About the Effects of Augsburg Forest). This work presents and discusses results in the context of the research project ForDigitHealth. The project is part of the Bavarian Research Association on Healthy Use of Digital Technologies and Media (ForDigitHealth), which is funded by the Bavarian Ministry of Science and Arts.

## Authors' Contributions

MB and YSC together conceptualized and designed the paper and interpreted the results. MB performed coding, data processing, and analysis. EA supervised the study, provided valuable suggestions for improvement, and critically revised the manuscript.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Three different scenarios for adding noise with differential privacy: adding noise to only shared layers, task-specific layers, and full layers.

[\[PNG File \(Portable Network Graphics File\), 30 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

Stress versus identity recognition performance of multitask centralized learning using 2 different datasets.

[\[PNG File \(Portable Network Graphics File\), 53 KB-Multimedia Appendix 2\]](#)

## References

1. De Silva LC, Miyasato T, Nakatsu R. Facial emotion recognition using multi-modal information. Presented at: Proceedings of ICICS, International Conference on Information, Communications and Signal Processing Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications; Sep 12, 1997:397-401; Singapore. [doi: [10.1109/ICICS.1997.647126](https://doi.org/10.1109/ICICS.1997.647126)]
2. Schuller B, Rigoll G, Lang M. Hidden Markov model-based speech emotion recognition. Presented at: International Conference on Acoustics, Speech and Signal Processing (ICASSP'03); Jul 6-9, 2003:II-1; Hong Kong, China. [doi: [10.1109/ICASSP.2003.1202279](https://doi.org/10.1109/ICASSP.2003.1202279)]
3. Wearables unit shipments worldwide by vendor from 2014 to 2022 (in millions), by vendor. Statista. 2020. URL: <https://www.statista.com/statistics/515634/wearables-shipments-worldwide-by-vendor/> [Accessed 2023-02-25]
4. Zhang C, Hu X, Xie Y, Gong M, Yu B. A privacy-preserving multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. Front Neurorobot. 2019;13:112. [doi: [10.3389/fnbot.2019.00112](https://doi.org/10.3389/fnbot.2019.00112)] [Medline: [31992979](https://pubmed.ncbi.nlm.nih.gov/31992979/)]

5. Zhang Y, Yang Q. A survey on multi-task learning. *IEEE Trans Knowl Data Eng.* 2022;34(12):5586-5609. [doi: [10.1109/TKDE.2021.3070203](https://doi.org/10.1109/TKDE.2021.3070203)]
6. Dai W, Cahyawijaya S, Bang Y, Fung P. Weakly-supervised multi-task learning for multimodal affect recognition. arXiv. Preprint posted online on Apr 23, 2021. [doi: [10.48550/arXiv.2104.11560](https://doi.org/10.48550/arXiv.2104.11560)]
7. Sang DV, Cuong LTB, Van Thieu V. Multi-task learning for smile detection, emotion recognition and gender classification. Presented at: SoICT '17: Proceedings of the 8th International Symposium on Information and Communication Technology; Dec 7-8, 2017:340-347; Nha Trang City, Viet Nam. [doi: [10.1145/3155133.3155207](https://doi.org/10.1145/3155133.3155207)]
8. Connolly HL, Young AW, Lewis GJ. Recognition of facial expression and identity in part reflects a common ability, independent of general intelligence and visual short-term memory. *Cogn Emot.* Sep 2019;33(6):1119-1128. [doi: [10.1080/02699931.2018.1535425](https://doi.org/10.1080/02699931.2018.1535425)] [Medline: [30336725](https://pubmed.ncbi.nlm.nih.gov/30336725/)]
9. Sheng W, Li X. Multi-task learning for gait-based identity recognition and emotion recognition using attention enhanced temporal graph convolutional network. *Pattern Recognit DAGM.* Jun 2021;114:107868. [doi: [10.1016/j.patcog.2021.107868](https://doi.org/10.1016/j.patcog.2021.107868)]
10. McMahan B, Moore E, Ramage D, Hampson S, y Arcas BA. Communication-efficient learning of deep networks from decentralized data. Presented at: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics; Apr 20-22, 2017:1273-1282; Fort Lauderdale, FL, United States.
11. Abadi M, Chu A, Goodfellow I, et al. Deep learning with differential privacy. Presented at: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security; Oct 24-28, 2016:308-318; Vienna, Austria. [doi: [10.1145/2976749.2978318](https://doi.org/10.1145/2976749.2978318)]
12. Regulation (EU) 2016/679 of the European Parliament and of the Council. Regulation (EU), 679. European Union; 2016. URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj> [Accessed 2024-11-11]
13. de la Torre L. A guide to the California Consumer Privacy Act of 2018. SSRN. Preprint posted online on Dec 21, 2018. [doi: [10.2139/ssrn.3275571](https://doi.org/10.2139/ssrn.3275571)]
14. Chen S, Jin Q, Zhao J, Wang S. Multimodal multi-task learning for dimensional and continuous emotion recognition. Presented at: AVEC '17: Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge; Oct 23, 2017:19-26; Mountain View, CA, United States. [doi: [10.1145/3133944.3133949](https://doi.org/10.1145/3133944.3133949)]
15. Can YS, Ersoy C. Privacy-preserving federated deep learning for wearable IoT-based biomedical monitoring. *ACM Trans Internet Technol.* Feb 28, 2021;21(1):1-17. [doi: [10.1145/3428152](https://doi.org/10.1145/3428152)]
16. Somandepalli K, Qi H, Eoff B, et al. Federated learning for affective computing tasks. Presented at: 2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII); Oct 18-21, 2022:1-8; Nara, Japan. [doi: [10.1109/ACII55700.2022.9953856](https://doi.org/10.1109/ACII55700.2022.9953856)]
17. Latif S, Khalifa S, Rana R, Jurdak R. Poster abstract: federated learning for speech emotion recognition applications. Presented at: 2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN); Apr 21-24, 2020:341-342; Sydney, NSW, Australia. [doi: [10.1109/IPSND48710.2020.00-16](https://doi.org/10.1109/IPSND48710.2020.00-16)]
18. Wang Z, Yang G, Dai H, Bai Y. DAFL: domain adaptation-based federated learning for privacy-preserving biometric recognition. *Fut Gen Comput Syst.* Jan 2024;150:436-450. [doi: [10.1016/j.future.2023.09.020](https://doi.org/10.1016/j.future.2023.09.020)]
19. Chhikara P, Singh P, Tekchandani R, Kumar N, Guizani M. Federated learning meets human emotions: a decentralized framework for human-computer interaction for IoT applications. *IEEE Int Things J.* 2020;8(8):6949-6962. [doi: [10.1109/JIOT.2020.3037207](https://doi.org/10.1109/JIOT.2020.3037207)]
20. Nandi A, Xhafa F. A federated learning method for real-time emotion state classification from multi-modal streaming. *Methods.* Aug 2022;204:340-347. [doi: [10.1016/j.ymeth.2022.03.005](https://doi.org/10.1016/j.ymeth.2022.03.005)] [Medline: [35314343](https://pubmed.ncbi.nlm.nih.gov/35314343/)]
21. Zhao Y, Chen J, Wu D, Teng J, Yu S. Multi-task network anomaly detection using federated learning. Presented at: SoICT '19: Proceedings of the 10th International Symposium on Information and Communication Technology; Dec 4-6, 2019:273-279; Hanoi, Ha Long Bay, Viet Nam. [doi: [10.1145/3368926.3369705](https://doi.org/10.1145/3368926.3369705)]
22. Smith V, Chiang CK, Sanjabi M, Talwalkar AS. Federated MTL. Presented at: Advances in Neural Information Processing Systems; Dec 4-9, 2017; Long Beach, CA, United States.
23. Shome D, Kar T. FedAffect: few-shot federated learning for facial expression recognition. Presented at: Proceedings of the IEEE/CVF International Conference on Computer Vision; Oct 11-17, 2021:4168-4175; Montreal, BC, Canada. [doi: [10.1109/ICCVW54120.2021.00463](https://doi.org/10.1109/ICCVW54120.2021.00463)]
24. Feng T, Peri R, Narayanan S. User-level differential privacy against attribute inference attack of speech emotion recognition on federated learning. arXiv. Preprint posted online on May 17, 2022. [doi: [10.48550/arXiv.2204.02500](https://doi.org/10.48550/arXiv.2204.02500)]
25. Dwork C, Roth A. The algorithmic foundations of differential privacy. *FNT Theor Comput Sci.* 2014;9(3-4):211-407. [doi: [10.1561/0400000042](https://doi.org/10.1561/0400000042)]
26. Wei K, Li J, Ding M, et al. User-level privacy-preserving federated learning: analysis and performance optimization. *IEEE Trans Mobile Comput.* 2021;21(9):3388-3401. [doi: [10.1109/TMC.2021.3056991](https://doi.org/10.1109/TMC.2021.3056991)]

27. Schmidt P, Reiss A, Duerichen R, Marberger C, Van Laerhoven K. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. Presented at: Proceedings of the International Conference on Multimodal Interaction; Oct 16-20, 2018; Boulder, CO, United States. [doi: [10.1145/3242969.3242985](https://doi.org/10.1145/3242969.3242985)]
28. Yadav M, Sakib MN, Nirjhar EH, Feng K, Behzadan AH, Chaspari T. Exploring individual differences of public speaking anxiety in real-life and virtual presentations. *IEEE Trans Affective Comput.* Jul 1, 2022;13(3):1168-1182. [doi: [10.1109/TAFFC.2020.3048299](https://doi.org/10.1109/TAFFC.2020.3048299)]
29. Lee Y, Lee N, Pham V, Lee J, Chung TM. Privacy preserving stress detection system using physiological data from wearable device. Presented at: Intelligent Human Systems Integration (IHSI 2023) Integrating People and Intelligent Systems; Feb 24-25, 2023; Venice, Italy. [doi: [10.54941/ahfe1002853](https://doi.org/10.54941/ahfe1002853)]
30. Rezaei S, Liu X. Multitask learning for network traffic classification. Presented at: 2020 29th International Conference on Computer Communications and Networks (ICCCN); Aug 3-6, 2020; Honolulu, HI, United States. [doi: [10.1109/ICCCN49398.2020.9209652](https://doi.org/10.1109/ICCCN49398.2020.9209652)]
31. Abd-Alrazaq A, Alajlani M, Ahmad R, et al. The performance of wearable AI in detecting stress among students: systematic review and meta-analysis. *J Med Internet Res.* Jan 31, 2024;26:e52622. [doi: [10.2196/52622](https://doi.org/10.2196/52622)] [Medline: [38294846](https://pubmed.ncbi.nlm.nih.gov/38294846/)]
32. Kargarandehkordi A, Slade C, Washington P. Personalized AI-driven real-time models to predict stress-induced blood pressure spikes using wearable devices: proposal for a prospective cohort study. *JMIR Res Protoc.* Mar 25, 2024;13(1):e55615. [doi: [10.2196/55615](https://doi.org/10.2196/55615)] [Medline: [38526539](https://pubmed.ncbi.nlm.nih.gov/38526539/)]
33. Li J, Washington P. A comparison of personalized and generalized approaches to emotion recognition using consumer wearable devices: machine learning study. *JMIR AI.* May 10, 2024;3(1):e52171. [doi: [10.2196/52171](https://doi.org/10.2196/52171)] [Medline: [38875573](https://pubmed.ncbi.nlm.nih.gov/38875573/)]
34. Wang M, Qin Y, Liu J, Li W. Identifying personal physiological data risks to the Internet of Everything: the case of facial data breach risks. *Humanit Soc Sci Commun.* 2023;10(1):216. [doi: [10.1057/s41599-023-01673-3](https://doi.org/10.1057/s41599-023-01673-3)] [Medline: [37192941](https://pubmed.ncbi.nlm.nih.gov/37192941/)]
35. El Rahman SA, Alluhaidan AS. Enhanced multimodal biometric recognition systems based on deep learning and traditional methods in smart environments. *PLoS One.* 2024;19(2):e0291084. [doi: [10.1371/journal.pone.0291084](https://doi.org/10.1371/journal.pone.0291084)] [Medline: [38358992](https://pubmed.ncbi.nlm.nih.gov/38358992/)]
36. Shen A, Francisco L, Sen S, Tewari A. Exploring the relationship between privacy and utility in mobile health: algorithm development and validation via simulations of federated learning, differential privacy, and external attacks. *J Med Internet Res.* Apr 20, 2023;25:e43664. [doi: [10.2196/43664](https://doi.org/10.2196/43664)] [Medline: [37079370](https://pubmed.ncbi.nlm.nih.gov/37079370/)]
37. Liu JC, Goetz J, Sen S, Tewari A. Learning from others without sacrificing privacy: simulation comparing centralized and federated machine learning on mobile health data. *JMIR Mhealth Uhealth.* Mar 30, 2021;9(3):e23728. [doi: [10.2196/23728](https://doi.org/10.2196/23728)] [Medline: [33783362](https://pubmed.ncbi.nlm.nih.gov/33783362/)]
38. Benouis M, Can YS, André E. A privacy-preserving multi-task learning framework for emotion and identity recognition from multimodal physiological signals. Presented at: 2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW); Sep 10-13, 2023:1-8; Cambridge, MA, United States. [doi: [10.1109/ACIIW59127.2023.10388160](https://doi.org/10.1109/ACIIW59127.2023.10388160)]
39. Almadhor A, Sampedro GA, Abisado M, et al. Wrist-based electrodermal activity monitoring for stress detection using federated learning. *Sensors (Basel).* Apr 14, 2023;23(8):3984. [doi: [10.3390/s23083984](https://doi.org/10.3390/s23083984)] [Medline: [37112323](https://pubmed.ncbi.nlm.nih.gov/37112323/)]
40. Fauzi MA, Yang B, Blobel B. Comparative analysis between individual, centralized, and federated learning for smartwatch based stress detection. *J Pers Med.* Sep 26, 2022;12(10):1584. [doi: [10.3390/jpm12101584](https://doi.org/10.3390/jpm12101584)] [Medline: [36294724](https://pubmed.ncbi.nlm.nih.gov/36294724/)]

## Abbreviations

- CL:** centralized learning
- CNN:** convolutional neural network
- DP:** differential privacy
- ECG:** electrocardiogram
- EDA:** electrodermal activity
- FedAvg:** federated averaging
- FL:** federated learning
- IID:** independent and identically distributed
- MFL:** multitask federated learning
- ML:** machine learning
- MTL:** multitask learning
- SGD:** stochastic gradient descent
- VERBIO:** virtual environment for real-time biometric interaction and observation

**WESAD:** wearable stress and affect detection

*Edited by John Torous; peer-reviewed by Eman Jaafar, Peter Washington, Puneet Gupta; submitted 28.04.2024; final revised version received 21.08.2024; accepted 22.08.2024; published 23.12.2024*

*Please cite as:*

*Benouis M, Andre E, Can YS*

*Balancing Between Privacy and Utility for Affect Recognition Using Multitask Learning in Differential Privacy-Added Federated Learning Settings: Quantitative Study*

*JMIR Ment Health 2024;11:e60003*

*URL: <https://mental.jmir.org/2024/11/e60003>*

*doi: [10.2196/60003](https://doi.org/10.2196/60003)*

© Mohamed Benouis, Elisabeth Andre, Yekta Said Can. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 23.12.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.