EAI Endorsed Transactions
on Pervasive Health and Technology
Research Article EAI.EU

Privacy-Preserving Human Motion Analysis for Lower Back Pain Stratification through Federated Learning

L. Gong[1], M. Yu[2,*] and S. P. Ardakani[2]

[1]School of Agri-food Technology and Manufacturing, University of Lincoln, Lincoln, UK
[2]School of Engineering and Physical Sciences, University of Lincoln, Lincoln, UK

Abstract

Human Gait Analysis is crucial in healthcare applications, with numerous research works focusing on machine learning and deep learning approaches for tasks such as abnormal gait detection and gait quality assessment. However, developing such models requires collecting and sharing a significant amount of patient data, raising privacy concerns. In this study, we introduce the world's first technique for constructing a deep neural network model to stratify patients' pain levels based on video recordings of timed up-and-go activities, while ensuring privacy preservation through modern federated learning algorithms. Our experimental results demonstrate the effectiveness of this technique in accurately stratifying LBP levels without the need for data sharing among local clients to maintain privacy.

Keywords: Federated Learning, Attention Model, Gait Analysis, Healthcare

Received on 28 August 2024, accepted on 01 November 2024, published on 17 April 2025

doi: 10.4108/eetpht.11.9109

[*]Corresponding author. Email: MYu@lincoln.ac.uk

1. Introduction

Human motion is intricately influenced by the coordinated function of the nervous, musculoskeletal, and cardiorespiratory systems, each playing a vital role in orchestrating fluid locomotion. Abnormal motions serve as notable indicators of underlying neurological, musculoskeletal, or biomechanical irregularities [1]. Therefore, effective human motion analysis holds profound significance, to early diagnosis and intervention across a spectrum of medical conditions impacting human mobility. Among the diverse applications of human motion analysis in healthcare, the stratification of lower back pain (LBP) severity emerges as a critical pursuit. LBP profoundly impacts individuals' quality of life and poses a significant societal burden. In the United Kingdom alone, back pain ranks as the largest single cause of disability, with lower back pain accounting for 11% of the total disability within the population [2]. Globally, an estimated 50– 80% of individuals are projected to experience challenges related to LBP, with a substantial 90% likely to undergo intermittent episodes even after initial alleviation [3].

Traditionally, questionnaire-based methods, such as the STarT Back Screening Tool (SBST) [4], have been commonly used to stratify the level of lower back pain (LBP). However, these approaches are subjective and rely on the clinician's expertise in assessing the severity of a patient's LBP. To enhance objectivity and precision in categorizing LBP severity, researchers have developed instrumentation-based methods. Innovative artificial intelligence (AI) algorithms utilize motion data collected from sensors to improve the classification of LBP. In [5], inertial measurement units (IMUs) attached to patients' trunks capture kinematic data related to trunk motion. This data is then processed to extract essential features such as

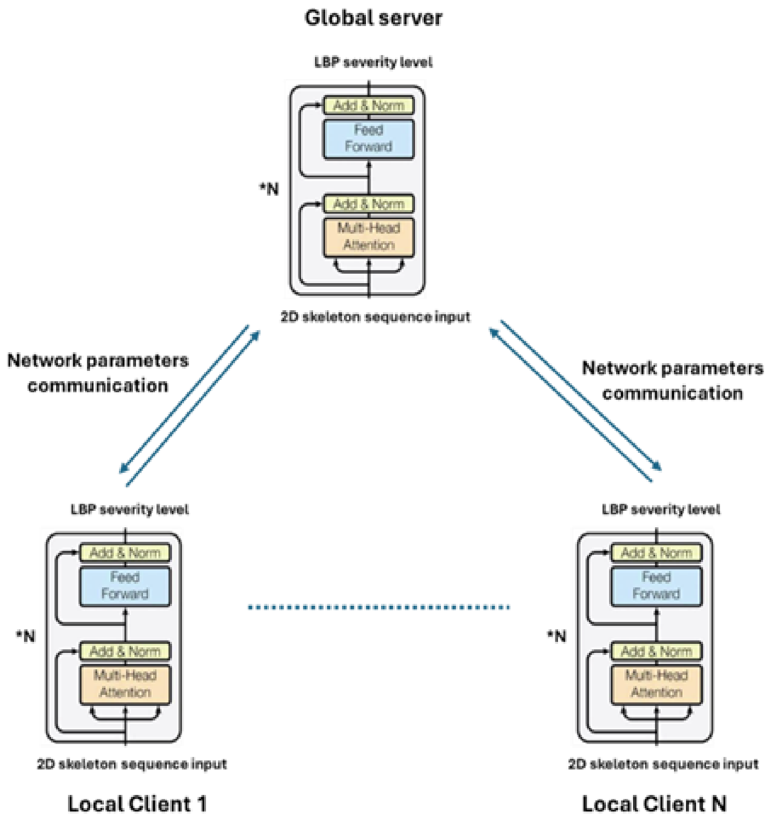

Figure 1. The diagram of the proposed federated learning-based technique for LBP severity stratification

linear acceleration and angular velocity. Machine learning algorithms like Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) are utilized to stratify LBP levels based on these features, with the SVM-based classifier demonstrating superior performance over MLP. The research work [6] involves a wireless inertial sensor network that collects signals from sensor units placed on various body regions. These signals are refined to derive parameters such as joint angles, providing insights into Lower Back Motion Assessment. In [7], researchers utilize IMU wearable sensors to capture motion data from a series of repetitive movements. Feature selection is performed to extract 25 variables, which are subsequently utilized to train multiple machine-learning models, which include logistic regression, decision tree, random forest, SVM, k-nearest neighbor (KNN), MLP, and gradient boosting algorithms for classifying subjects based on normal/abnormal movements. Experimental results showcase that SVM, random forest, and MLP models achieve classification accuracies exceeding 90%. By using unsupervised machine learning techniques on full-body biomechanics, encompassing kinematics, dynamics, and muscle forces, captured via marker-less Kinect depth cameras, [8] identifies a forward-leaning sit-to-stand strategy (STS) as a distinctive movement biomarker for LBP subjects. In [9], a more intricate 3D motion capture system is employed to analyze differences in spinal kinematics and 3D kinematic patterns between healthy individuals and those experiencing LBP.

To develop a sophisticated machine learning or deep learning model for stratifying low back pain (LBP) based on sensor recordings of human motions, a significant amount of patient motion data must be collected and shared from various sources. However, this raises privacy concerns, particularly when the motion data is captured using video cameras. To address this issue, we propose a novel technique in this study: training an attention-based deep neural network model [10] using a federated learning algorithm [11]. With this approach, the DNN model is trained on individual partitions of local motion data captured by a normal video camera, and the training results are then aggregated to create the final model. This method allows each part of the local data to remain on the local machines for training purposes without the need to be shared with other parties, thus preserving privacy. To the best of our knowledge, this is the first instance of federated learning being utilized for healthcare-related human motion analysis tasks.

2. Methodology

The proposed federated learning-based LBP stratification technique is illustrated in Figure 1. A cost-effective RGB camera sensor is utilized to capture video recordings of timed up-and-go (TUG) activities, which are stored on individual local clients. The Detectron2 library [12] is employed to extract 2D human poses from the recorded videos, which are subsequently pre-processed to ensure scale and position invariance as per [13]. An attention-based model is then implemented to stratify the severity level of LBP based on the preprocessed 2D human pose data.

As shown in Fig. 1, each attention model contains multiple blocks while each block contains multi-head attention, additional&normalization and feed-forward layers, which generate the output of the LBP severity level classification based on input 2D skeletons. The pivotal element within the attention model is the multi-head attention layer [10], which leverages multiple 'heads' to generate outputs. For the i-th head, queries (Q_i), keys (K_i) and values (V) are computed as $Q_i = XW_{Qi}$, $K_i = XW_{Ki}$ and $V_i = XW_{Vi}$ respectively, where X represents the input of the multi-head attention layer while W_{Qi}, W_{Ki} and W_{Vi} denote the respective parameter matrices associated with the i-th head. Based on Q_i, K_i and V_i, the attention weights A_i for the i-th head are calculated as:

$$A_i = softmax\left(\frac{Q_i K_i^T}{\sqrt{Dh}}\right) \qquad (1)$$

where D^h is a scale factor and $softmax(\cdot)$ is an activation function and the output of the i-th head denoted as H_i is calculated as:

$$H_i = A_i V_i \qquad (2)$$

which is the weighted summation of V_i based on the attention weights. All head outputs are calculated in the same way and finally concatenated and linearly projected as the final output of the multi-head attention layer as:

$$MSA(X) = [H_1;H_2;\ldots;H_N]W_{MSA} \qquad (3)$$

where $MSA(X)$ represents the multi-head attention layer output based on the input X, W_{MSA} is a projection matric and N is the head number in the multi-head attention layer. The output of a multi-head attention layer will then go through a series of add&norm operations and a small feed-forward network to generate the output of a block. The attention model parameters at the global server are updated by aggregating the training results from multiple local clients. In this study, we utilize the FedAvg algorithm [11] to facilitate this process, as outlined in Table 1. By leveraging the FedAvg algorithm, the model on the global server can be trained without necessitating the sharing or collection of data from all local clients. This approach ensures that data can remain secure within the local clients without the requirement to disclose it to a third party. Only the trained network parameters, rather than the original data, are transmitted for updating the global network architecture. This methodology effectively safeguards data privacy.

Table 1. The description of the FedAvg Algorithm

Algorithm: Federated Averaging:

Server executes:
Initialize the network parameters w0
for each round t=1,…, do
for each client k:

$$w_{t+1}^k \leftarrow \text{ClientUpdate}\ (w_t^k)$$
$$w_{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w_{t+1}^k$$

ClientUpdate(w):
for each local epoch I from 1 to E do
for batch b in batches do
update local network weight **w**
Return **w** to the server

3. Experimental studies

We evaluated the proposed technique using a dataset obtained through collaboration with a physiotherapy clinic in Lincoln, UK. The dataset consists of video recordings of TUG activities performed by 21 subjects, including 9 males and 12 females, and classified into low/medium/high LBP severity categories by specialized physiotherapists. The acquired video recordings were divided into 30-frame video clips using the sliding window technique with a sliding length of 5 frames. Each video clip was analyzed using the Detectron 2 library to extract 2D skeleton sequences from the original video frames. Figure 2 depicts the extracted 2D skeletons from a TUG activity in one of the video clips.

We have created both training and testing datasets for our study. The training dataset is divided into three parts, each containing approximately 1,750 skeleton sequence samples, mirroring a scenario where data is distributed across three local clients. The testing dataset comprises 2,623 samples. We trained an attention model which include 9 blocks and 8 heads, using the FedAvg learning algorithm in Table 1 on the three partitions of the training dataset and tested it on the testing dataset. Concerning the

Figure 2. Samples of video frames and corresponding 2D skeleton extractions via Detectron2

network training on the local client, the AdamW algorithm [14] with specific hyperparameters: a learning rate of 0.0005, a weight decay rate of 0.01, and a learning rate decay of 0.99. The evaluation is conducted based on the following metrics:

$$precision = \frac{TP}{TP + FP} \tag{4}$$

$$recall = \frac{TP}{TP + FN} \tag{5}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$f1_score = \frac{2 \cdot precision \cdot recall}{precision + recall} \tag{7}$$

TP, TN, FP, and FN denote true positives (correctly classified class samples), true negatives (correctly classified non-class samples), false positives (incorrectly classified class samples), and false negatives (incorrectly classified non-class samples) for a given class. The closer these values are to 1, the more indicative they are of a model's effectiveness.

Table 2. Comparison between FedAvg and training on local datasets

	accuracy	precision	recall	F1_score
Partition one	91.92%	92.27%	91.92%	91.97%
Partition two	68.62%	68.60%	68.62%	67.29%
Partition three	82.81%	82.88%	82.81%	82.82%
FedAvg	**95.04%**	**95.25%**	**95.04%**	**95.04%**

The results are presented in Table 2. Additionally, we provide results obtained from training on a single partition of local data for comparison. The findings indicate that the FedAvg algorithm, which integrates multiple data partitions, outperforms using only one local data partition. Furthermore, we conducted a comparison of the attention-based model with other traditional deep learning models (e.g., LSTM, CNN) and showcased the results in Table 3. It is evident from the comparison that the proposed approach demonstrates superior performance compared to other models.

Table 3. Comparison between different models

	accuracy	precision	recall	F1_score
LSTM	62.26%	68.84%	62.26%	60.13%
CNN	90.55%	91.38%	90.55%	90.33%
Attention model	**95.04%**	**95.25%**	**95.04%**	**95.04%**

4. Conclusions

In this study, we introduce a novel federated learning-based attention model for stratifying LBP severity using video recordings of TUG activities. Specifically, we employ the FedAvg algorithm to aggregate training results from local datasets on multiple clients and update the attention model on the global server. Our experimental findings demonstrate that the proposed approach achieves high accuracy in classifying LBP severity. Additionally, privacy concerns are addressed through the federated learning algorithm, which only communicates network parameters for model training without the need to collect datasets from multiple local clients. Moving forward, we plan to explore various network architectures and federated learning algorithms to further enhance performance.

Acknowledgements.

This research has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 778602 ULTRACEPT.

References

[1] Pirker, W, Katzenschlager, R.: Gait disorders in adults and the elderly: A clinical guide. Wien Klin Wochenschr, 129 (3), 81-95 (2017)

[2] The National Back Pain Pathway, NHS England, https://www.england.nhs.uk/blog/charlesgreenough/, last accessed 2024/07/11.

[3] Fatoye, F., Gebrye, T., Odeyemi, I.: Real-world incidence and prevalence of low back pain using routinely collected data, Rheumatol. Int., 39(4), 619–626, (2019).

[4] Hill, C., Dunn, K., Lewis, M., Mullis, R., Main, C., Foster, N., Hay, E.: A primary care back pain screening tool: Identifying patient subgroups for initial treatment, Arthritis Care Res.: Off. J. Am. Coll. Rheumatol., 59, 632-641 (2008).

[5] Abdollahi, M., Ashouri, S., Abedi, M., Azadeh-Fard, N., Parnianpour, M., Khalaf, K., Rashedi, E.: Using a Motion Sensor to Categorize Nonspecific Low Back Pain Patients: A Machine Learning Approach, Sensors, 20(2), (2020).

[6] Molnar, M., Kok, M., Engel, T., Kaplick, H., Mayer, F., Seel, T.: A Method for Lower Back Motion Assessment Using Wearable 6D Inertial Sensors, 21st International Conference on Information Fusion (FUSION), Cambridge, UK (2018).

[7] Villalba-Meneses, F. et al. : Classification of the Pathological Range of Motion in Low Back Pain Using Wearable Sensors and Machine Learning, Sensors, 24(3), (2024).

[8] Keller, A., Torres-Espin, A., Peterson, T., Booker, J., O'Neil, C., Lotz, J., Bailey, J., Ferguson, A., Matthew, R.: Unsupervised Machine Learning on Motion Capture Data Uncovers Movement Strategies in Low Back Pain, Frontiers in Bioengineering and Biotechnology, 10, (2022).

[9] Hemming, R., Sheeran, L., Deursen R., Sparkes, V.: Non-specific chronic low back pain: differences in spinal kinematics in subgroups during functional tasks, Eur Spine J., 27(1), 163-170 (2018).

[10] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., Polosukhin, I.: Attention Is All You Need, 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA (2017).

[11] McMahan, H., Moore, E., Ramage, D., Hampson, S., Arcas, B.: Communication-Efficient Learning of Deep Networks from Decentralized Data, https://arxiv.org/abs/1602.05629, last accessed 2024/07/11.

[12] Wu, Y., Kirillov, A., Massa, F., Lo, WY., Girshick, R.: Detectron2, https://github.com/facebookresearch/detectron2, last accessed 2024/07/11.

[13] McCall, S., Kolawole, S., Naz, A., Gong, L., Ahmed, S., Prasad, P., Yu, M., Wingate, J., Ardakani, S.: Computer Vision Based Transfer Learning-Aided Transformer Model for Fall Detection and Prediction, IEEE Access, 12 (2024).

[14] I. Loshchilov, I., Hutter, F.: Decoupled Weight Decay Regularization, International Conference on Learning Representations, Louisiana, United States (2019).