TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON)
31 Oct - 3 Nov 2023. Chiang Mai, Thailand

A Combination of Feature Extraction and Feedforward Neural Network to Estimate Muscle Activity in Human Gait

Min Khant[1], Darwin Gouwanda[1]*, Alpha A. Gopalai[1], King Hann Lim[2], and Chee Choong Foong[3]

[1] School of Engineering, Monash University Malaysia, Bandar Sunway, Malaysia;
min.khant@monash.edu; darwin.gouwanda@monash.edu; alpha.agape@monash.edu
[2] Department of Electrical & Computer Engineering, Curtin University, Miri, Malaysia;
glkhann@curtin.edu.my
[3] Sunway Medical Centre, Bandar Sunway, Malaysia; foongcc@sunway.com.my
* Correspondence: darwin.gouwanda@monash.edu;

***Abstract*—Inertial Measurement Unit (IMU) has been widely recognized to be the practical alternative to capture and analyze human gait. However, due to its inherent characteristics, it can only measure the basic kinematics of the body segment it attached to. With the help of the machine learning, IMU can be used to determine the dynamic behavior of the major lower extremity muscle. This paper explores the use of feature-extracted IMU data and a neural network to estimate muscle activity during walking. IMU and Electromyogram (EMG) data were collected from fifty-eight healthy participants. Principal Component Analysis (PCA) and Tsfresh (Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests) were applied to extract the relevant features from the data. These features were then used to train the Feedforward Neural Network (FNN). A combination of Tsfresh and FNN yielded the best results with correlation coefficient (*r*) of 95.73% and Root Mean Square Error (*RMSE*) of 11.20%. This research can potentially help reduce the number of sensors needed in gait analysis, allow for portable motion capture, and improve the accuracy and efficiency of the FNN model in estimating muscle activity.**

Keywords—Inertial Measurement Unit (IMU), Electromyogram (EMG), Feedforward Neural Network (FNN), Principal Component Analysis (PCA), Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (Tsfresh).

I. INTRODUCTION

The application of wearable sensors, such as Inertial Measurement Unit (IMU) has gained significant traction in the human movement analysis. This sensor provides critical information about the body motion in various activities, including walking. Typical walking gait analysis involves the use of optical motion capture system, force plate and Electromyogram (EMG). However, these systems are bulky and require elaborate setup. Thus, performing gait analysis outside of the clinical or laboratory setting is difficult.

The capacity to perform gait analysis beyond the confine of a controlled environment presents an opportunity to gather real-world data. This advantage holds great significance, particularly in monitoring patients with mobility impairment in their activities of daily living and in the comfort and familiarity of their own homes. The main characteristics of the IMU, which are light and inexpensive, and can be easily attached to various part of human body, offer a practical alternative to the conventional measurement techniques. It can help make gait analysis more accessible and open up novel applications and research, some of which can be seen in [1, 2].

With the right signal processing technique and machine learning, IMU has been reported to be able to extract various gait parameters and joint kinetics and kinematics [3, 4]. However, there is a lack of literature on predicting muscle activity using IMU and neural network. In the previous study [5], we attempted to estimate the EMG signal using a Feedforward Neural Network (FNN) and Long-Short Term Memory (LSTM) along with IMU data from an online open-access dataset. Other studies proposed the use of musculoskeletal modelling [6] and Feedforward nonlinear autoregressive model with exogenous (NARX) [7] to determine muscle behavior using joint kinetics and kinematics data obtained from motion capture system.

Principal Component Analysis (PCA) is commonly used in gait analysis to eliminate redundant information and to help understand multiple gait signals [8, 9]. PCA is a dimensionality reduction technique that projects data into a lower-dimensional linear subspace [10, 11]. It does this by finding the linear basis that captures the maximum variance in the data. Due to it being lightweight and an unsupervised algorithm (not needing a target to extract features), it is an ideal tool for real-time analysis and interpretation of gait data. It allows for quick and efficient assessment of gait performance in clinical settings. One of the clinical applications of this method includes using six dimensions IMU data and PCA to give cyclograms that clinicians can use to assess gait performance [12]. The study reported in [13] used trunk IMU and PCA to identify the limb affected by stroke.

The other feature extraction method worth exploring is the supervised feature extraction method called Tsfresh (Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests) [14]. This method calculates a large number of time series features and helps evaluate the explaining power and importance of the features for regression or classification tasks. Tsfresh on IMU and EMG

This study was supported by Malaysia Ministry of Higher Education under the Fundamental Research Grant Scheme (FRGS) with project Ref. No. FRGS/1/2020/TK0/MUSM/02/2.

data along with Random Forest (RF) regressor were used to estimate joint kinetics and kinematics and muscle force [15]. Studies reported in [16, 17] demonstrated the combination of mobile phone accelerometer, Tsfresh and neural network to classify different activities such as walking, running, sitting, jumping and more. Another study applied Tsfresh to extract features from pelvis IMU data and identified abnormal gait using decision-tree-based learning algorithms [18].

This paper aims to investigate the effects of different feature extraction techniques i.e. PCA and Tsfresh on FNN in estimating lower extremity muscle activity in gait. PCA is selected because it is a lightweight and unsupervised algorithm that can extract features quickly without the intended target. On the contrary, Tsfresh is a comprehensive supervised feature extraction method. It is believed a combination of feature extraction and FNN can yield good estimation results, which can potentially improve the accuracy and efficiency of gait analysis and provide valuable insights into the relationship between gait IMU data and EMG data. In addition, this method can reduce the number of devices attached to lower extremity, making it more convenient for gait analysis. By using only IMU, gait analysis can be performed anywhere outside of the traditional laboratory, providing flexibility and opportunity to study gait in real-world settings.

II. Methods

A. Data Collection

The gait data used in this study is collected from fifty-eight healthy adults (26 males, 19-73 years old, 64.34 ± 19.51 kg, 160.29 ± 17.7 cm). The participant selection criteria were adults over 18 years old, without any lower limb injury to the lower limb in the past six months and able to walk comfortably and without altered gait. The process was carefully explained to the participants, and each participant was asked to sign an informed consent. The study was approved by the Monash University Human Research Ethics Committee (Project number: 32328).

First, eight sensors (six Delsys Trigno and two Delsys Duo, Massachusetts, USA), which provided eight channels of both IMU and EMG data, were attached to the participants. The EMG from Delsys Trigno and DUO were sampled at 1,259 Hz and 1,778 Hz, respectively. As for the IMU, they were sampled at 148 Hz by Delsys Trigno sensor and 963 Hz by the accelerometer and 741 Hz by the gyroscope in the Delsys Duo sensor. The sampling frequencies are predetermined and fixed with no means of altering or adjusting these frequencies. The EMG electrodes were used to record the muscle activity of Tibialis Anterior, Gastrocnemius Medialis, Gastrocnemius Lateralis, Rectus Femoris, Vastus Lateralis, Vastus Medialis, Semitendinosus and Bicep Femoris. The selection of these particular muscles is based on their prevalence in existing literature, suitability for surface EMG, and they cover major segments of the lower extremity used in gait. The sensors were attached according to SENIAM standards [19]. Of the eight available sensors of IMU, only four sensors that were attached to foot, Tibialis Anterior, Rectus Femoris and trunk, were used in this work. Each IMU has six channels: three axes acceleration and three axes angular velocity. This results in a total of 24 channels. The sensor placement and its sensing axes are shown in Fig. 1.

The participants were asked to walk for five trials at three different speeds (slow, normal, and fast) on the platform with three force plates (Bertec, Ohio, USA) sampled at 1,000 Hz as shown in Fig. 2. In each trial, the participants would start walking from one end to the other end, then turn around and walk back to the other end while stepping on one force plate at a time. A total 1,740 gaits were collected.

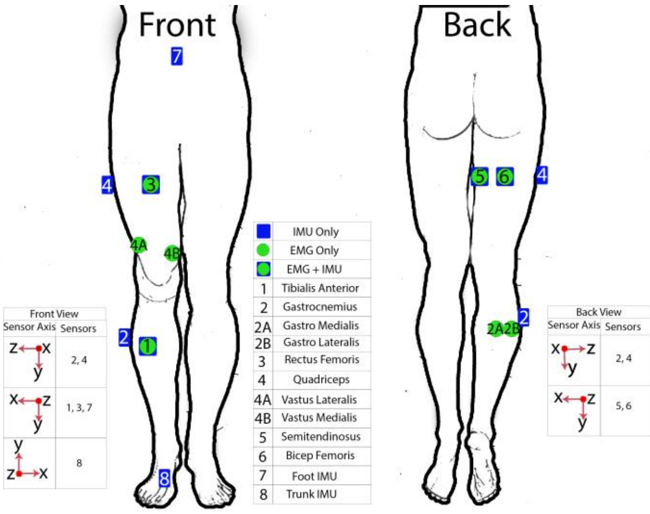


Fig. 1. Sensor positions

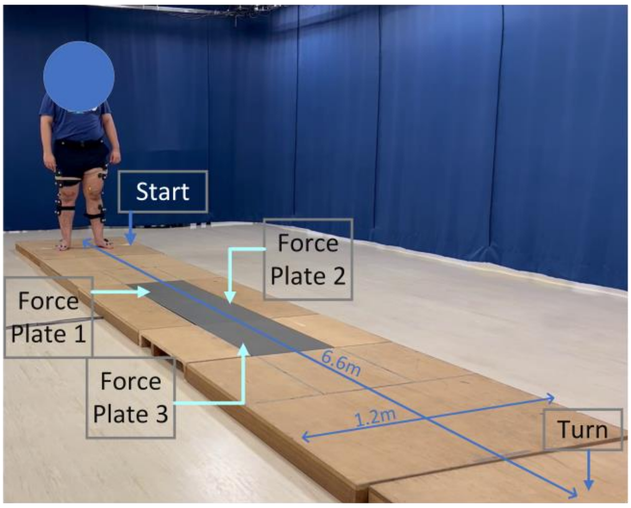


Fig. 2. The walkway for participants to walk on

B. Data processing

The vertical ground reaction force exerted by the foot during walking was used to identify the heel-strike and to segment the gait. Each gait cycle was visually inspected, and only those with dominant leg heel strikes on the first force plate were selected. This resulted in a total of 848 gaits. The force plate data were then filtered using a Butterworth lowpass filter with cut-off frequency of 10 Hz. The time of heel strike was determined by selecting a threshold of 10 N, as suggested in [20].

The EMG and IMU data from Delsys Duo were down sampled to 1259 Hz and 148 Hz, respectively. The data were then segmented per gait cycle using the timing of heel strike recorded by Force Plate 1 and Force Plate 3. Next, the segmented EMG data were filtered using a Butterworth band pass filter (4^{th} order, 20 Hz, 400 Hz) to reduce noise. The filtered data were then rectified and passed through a Butterworth lowpass filter (4^{th} order, 4 Hz) to get the enveloped EMG data. The data were time-normalized to 101 data points, representing the percentage of the gait cycle. They were then filtered using a median filter to reduce the impulse noise, and min-max normalized to values ranging between 0 and 1. Similar to the EMG, the IMU data were time-normalized to 101 data points. A median filter was applied, followed by min-max normalization.

Three subjects from different age groups (19-29, 30-49, 50-73 years old) were randomly selected to be used as unseen test data. These data have 44 gait cycles. Afterwards, the remaining data were divided into 3 datasets (training, validation, and testing) with a ratio of 80:15:5. This means that the dataset was made up of 642 gaits training data, 120 gaits validation data, 41 gaits testing data and 44 gaits unseen test data.

C. Feature Extraction

PCA is a technique that simplifies a dataset by reducing the number of variables while retaining as much statistical information as possible. This is achieved by finding new variables called Principal Components (PCs) through solving an eigenvalue or eigenvector problem [21]. First, the mean of input along variable j ($\overline{X}_j$) was calculated, where j is the IMU channel. Then the difference between the input (X_{ij}) and the mean was calculated to obtain the centered variable (X^*_{ij}) as shown in (1), where i is along the time. Next, covariance matrix (S) was calculated using (2) where X^* was the center variable and $X^{*\prime}$ was the transpose of the centered variable (X^*). The covariance matrix (S) obtained can be used to find a set of eigenvalues (λ) and eigenvectors (a) by solving (3). This leads to a set of eigenvectors and eigenvalues of the covariance matrix, with the largest eigenvalue representing the linear combination with the highest variance. The eigenvectors were sorted based on their corresponding eigenvalues in descending order and PC loading (a_k) was obtained by selecting the k number of sorted eigenvectors. Finally, principal components (Y) can be calculated using (4) to obtain a new dataset with k number of components. The representation or cumulative variance was calculated by summating all the explained variance ratios. This ratio represents how much variation in the data was captured by each component [22].

$$X^*_{ij} = X_{ij} - \overline{X}_j \tag{1}$$

$$(n\text{-}1)S = X^{*\prime}X^* \tag{2}$$

$$Sa = \lambda a \tag{3}$$

$$Y = X^* a_k \tag{4}$$

PCA was applied to each gait dataset that has a shape of 101 datapoints × 24 channels. It was found that the first 13 out of 24 components can represent 99% of the data. This results in data with an output shape of 101 data points × 13 features.

Tsfresh was applied to the same dataset to extract statistically significant features from the measurements collected by IMU. Tsfresh combines 63 time series characterization methods (such as median, maximum, minimum, absolute energy and Fast Fourier Transform (FFT)) to compute 783 features from each IMU channel in its default setting [14]. This means that 24 channels produced 18,792 features. Tsfresh also contains a feature selection method based on automatically configured hypothesis tests that allow for identifying statistically significant time series characteristics. By providing this feature selection method with extracted features and the required target (EMG), it identified 244 relevant features that are statistically significant. Afterwards, these features were ranked according to their importance to each target using Random Forest (RF) regressor. Random forest combines multiple decision trees and trains each tree with different features. It computes the feature importance of each feature by accumulating the error reduction contributed by each variable during the fitting process [23]. The top 20 most important features for each target were selected, and repeating features were removed to avoid redundancy. This resulted in 74 features, which were used to train and test the neural network.

D. Neural Network

Feedforward Neural Network (FNN) model was constructed in this study. The model consists of one input layer, one output layer and four hidden (dense fully connected) layers. Each hidden layer has 256 neurons and a dropout layer between each hidden layer. The desired output is a 1D array of normalized EMG data for each individual muscle. Three models are trained for each type of input: (1) Raw cascaded normalized 3D IMU data; (2) PCA features and (3) Tsfresh features. This is summarized in Fig. 3.

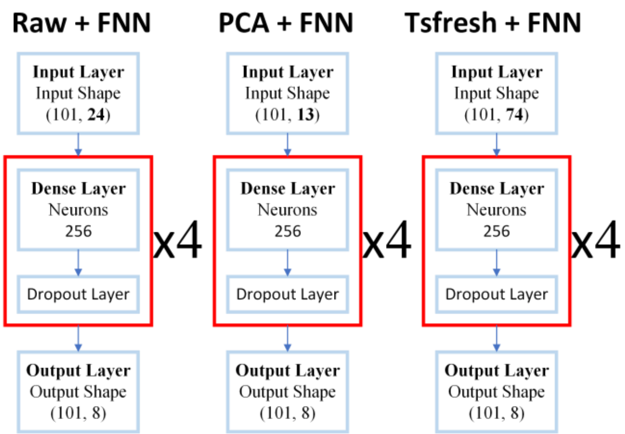


Fig. 3. Neural network architecture with 3 input shapes

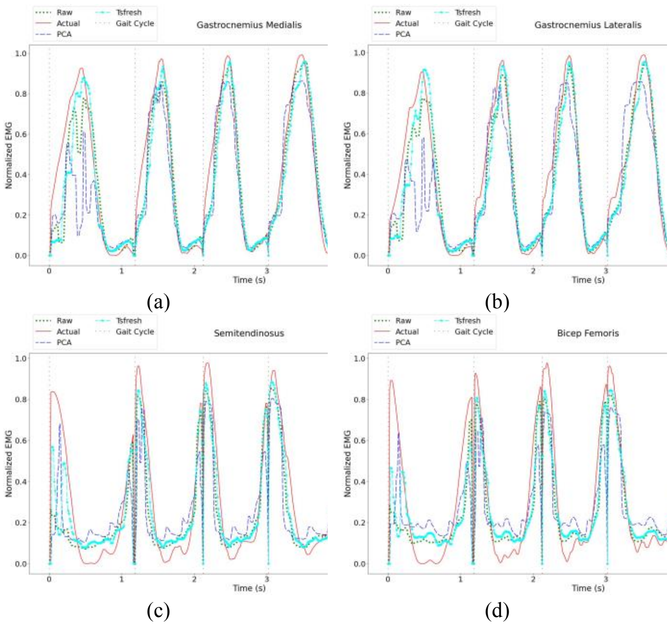
III. Results

The average of Root Mean Squared Error (*RMSE*) and correlation coefficient (r) of each muscle are presented in Table 1. Tsfresh produced the best results when using the test datasets. Tsfresh has *RMSE* less than 15% and r greater than 90%. Although PCA performed worse than Tsfresh, the results are still encouraging. It has *RMSE* lower than 20% and r higher than 80%. In the case of Gastrocnemius Medialis and Lateralis, *RMSE* is less than 15% and r greater than 90%. Using raw data as input gives better results than PCA. The *RMSE* was found to be lower than 15% and r was greater than 90% for all the investigated muscles. The only exception is the tibialis anterior, which has *RMSE* greater than 14% and r less than 90%. Fig. 4 shows a sample of the EMG waveform of two best (Gastrocnemius Medialis and Lateralis) and two worst estimations (Semitendinosus and Biceps Femoris).

Next, the unseen subjects' data were used to estimate muscle activities and the results are shown in Table 2. Tsfresh and raw data have similar results, with three of the muscles (Gastrocnemius Medialis, Lateralis, and Vastus Lateralis) having *RMSE* less than 15% and *r* greater than 90%. They performed poorly in estimating the muscle behavior of Vastus Medialis, Semitendinosus and Bicep Femoris. The *RMSE* are greater than 20% and *r* are lower than 75%. Similar to test dataset results, PCA performed the worst as most results have *RMSE* above 20% and *r* below 75%. Only 2 muscles (Gastrocnemius Medialis and Gastrocnemius Lateralis) have *RMSE* below 20% and *r* over 75%. A sample of the estimated muscle activities of Gastrocnemius Medialis, Gastrocnemius Lateralis, Semitendinosus and Biceps Femoris are presented in Fig. 5.

TABLE I. ACTUAL AND PREDICTED MUSCLE ACTIVITIES COMPARISON FOR TEST DATA

	Tsfresh+FNN		***PCA+FNN***		***Raw+FNN***	
	RMSE (%)	*r (%)*	*RMSE (%)*	*r (%)*	*RMSE (%)*	*r (%)*
Tibialis Anterior	13.56	89.11	16.80	82.50	14.94	86.47
Gastrocnemius Medialis	7.36	97.54	12.02	93.26	9.81	95.75
Gastrocnemius Lateralis	7.41	97.39	13.15	91.25	9.36	95.73
Rectus Femoris	9.12	96.11	14.74	87.75	9.70	94.89
Vastus Lateralis	8.01	96.90	15.23	87.54	9.84	95.01
Vastus Medialis	9.52	95.38	15.79	86.21	10.97	93.56
Semitendinosus	11.00	92.72	17.16	81.32	12.82	89.96
Bicep Femoris	11.65	91.74	17.23	81.42	12.27	90.71

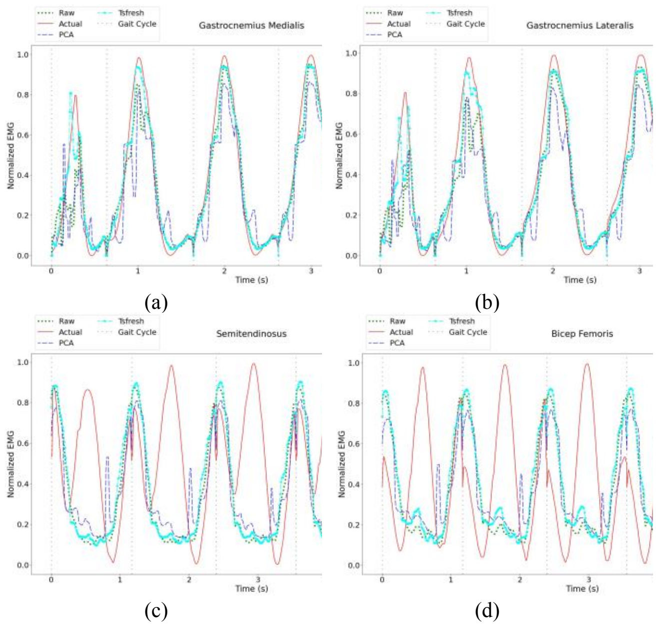


(a) (b) (c) (d)

Fig. 4. A sample of actual and predicted muscle activities for (a) Gastrocnemius Medialis (b) Gastrocnemius Lateralis (c) Semitendinosus (d) Bicep Femoris from testing dataset.

TABLE II. ACTUAL AND PREDICTED MUSCLE ACTIVITIES COMPARISON FOR UNSEEN DATA

	Tsfresh+FNN		***PCA+FNN***		***Raw+FNN***	
	RMSE (%)	*r (%)*	*RMSE (%)*	*r (%)*	*RMSE (%)*	*r (%)*
Tibialis Anterior	18.45	79.50	21.05	72.23	18.23	80.07
Gastrocnemius Medialis	11.20	95.73	17.48	89.11	13.65	92.98
Gastrocnemius Lateralis	12.11	93.76	18.22	86.77	13.11	93.10
Rectus Femoris	14.35	89.40	23.18	68.59	14.98	88.65
Vastus Lateralis	12.26	92.26	21.35	74.08	12.26	92.23
Vastus Medialis	21.66	73.90	23.83	66.91	21.30	75.02
Semitendinosus	25.68	63.64	25.98	61.34	26.11	62.94
Bicep Femoris	24.20	63.62	25.37	56.04	24.49	62.13



(a) (b) (c) (d)

Fig. 5. A sample of actual and predicted muscle activities for (a) Gastrocnemius Medialis (b) Gastrocnemius Lateralis (c) Semitendinosus (d) Bicep Femoris from unseen subject (age: 68 years old) data.

IV. DISCUSSION

Based on the results obtained from the unseen data, it is evident that most of the results from Tsfresh and raw data models have *r* above 80% and *RMSE* below 20%. A *r* value greater than 80% signifies a strong positive relationship between the actual and predicted values, indicating the model's ability to capture the overall trend in the data and provide a reasonable estimation of the target variable based on the input features. *RMSE* value below 20% denotes that, on average, the predicted values deviate from the actual values by less than 20% of the range of the target variable. This suggests that the model's prediction is relatively close to the true values and have low errors on unseen data. These results indicate that the model could be applicable in real-world scenarios, where accurate predictions are essential. All three methods performed well when the test data were used. However, PCA results were significantly affected when unseen data were used. PCA is an unsupervised feature extraction method, meaning it does not need any target data. This implies that the features extracted using this method might not be optimal. Another limitation of the PCA is that the underlying structure of the data must be linear, so it may not adequately capture the nonlinear patterns present in time series data. As time series data is sequential in nature and PCA treats all datapoints as independent [24], it may not capture the important temporal patterns. In addition, if there are multiple patterns in the data that are closely related to each other, PCA may not be able to distinguish them [22].

Using Tsfresh data with FNN is better than just using raw input. Although the features improved the estimation results, the process is computationally costly. A good processing unit is preferred to perform multiple statistical calculation. In this study, the process took about seven seconds for each gait when it was run on a machine with one Intel Xeon Gold 6338 core and 32 GB of RAM. Even though the time could be reduced by only calculating the required parameters, it would not be suitable for real time processing.

Several methods can be considered to overcome these limitations. One of them is to use an LSTM model in place of the FNN. As previously shown [5], LSTM performs much better than FNN for this task due to its ability to retain information for a long period of time. Another model that could be explored is CNN-LSTM [25]. In this model, CNN layers can act as feature extraction layer, and LSTM can be used to predict the muscle activity. In addition to using different machine learning architectures, the frequency-domain of the IMU data could be considered because the majority of features that Tsfresh provided have Fast-Fourier transformed data. Although Tsfresh is not suitable for real-time use, it has provided important attributes, such as FFT, that can be calculated without relying on this algorithm.

In a normal gait, some muscles such as Vastus Medialis, Vastus Lateralis, Biceps Femoris and Semitendinosus activate between terminal swing and opposite toe-off phase of the gait cycle [26]. Since the data were segmented using heel-strikes, these muscle activities occurred between two consecutive gait cycles – at the beginning and at the end of the gait cycle. This creates a rather odd and segmented waveform in every gait cycle as can be seen in Fig 4(c) and Fig. 4(d). Hence, the segmentation method might not be appropriate for these muscles, thus led to lower estimation results. Training the machine learning models using data segmented with both heel-strike and toe-off could potentially enhance their performance.

The experimental results (Fig. 5c and Fig. 5d) indicate that some elderlies activated their hamstring muscles i.e. Semitendinosus and Biceps Femoris twice during the stance phase. This behavior generated an additional peak muscle activity. This peak is typically absent in young adults, and it is not considered to be a part of the normal gait [26]. Nevertheless, similar observations were found and reported in [27]. It was speculated that these activities might be attributed to the stiffening of the knee joint, serving as a mechanism to enhance stability in the elderly.

V. Conclusion

This paper investigates the impacts of two distinct feature extraction methods i.e. PCA and Tsfresh on the prediction of muscle activity using FNN and IMU data. In overall, Tsfresh performed better than PCA. This is reasonable considering that PCA is an unsupervised method. On the other hand, Tsfresh, while not leading to a significant improvement in the model accuracy, proved to be valuable in identifying essential features that could be utilized in future research. The findings further indicate that the performance of the two hamstring muscles was comparatively suboptimal, potentially influenced by the variations in muscle activities between elderly and younger adults, which may also contribute to differences in stability. Future potential investigations could delve into alternative neural network architectures, such as LSTM, or explore alternative feature extraction methods. These endeavors could offer valuable insights and contribute to the advancement of the research domain.

References

[1] L. C. Benson, A. M. Räisänen, C. A. Clermont, and R. Ferber, "Is This the Real Life, or Is This Just Laboratory? A Scoping Review of IMU-Based Running Gait Analysis," *Sensors,* vol. 22, no. 5, p. 1722, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/5/1722.

[2] C. Monoli, J. F. Fuentez-Pérez, N. Cau, P. Capodaglio, M. Galli, and J. A. Tuhtan, "Land and Underwater Gait Analysis Using Wearable IMU," *IEEE Sensors Journal,* vol. 21, no. 9, pp. 11192-11202, 2021, doi: 10.1109/JSEN.2021.3061623.

[3] M. Sharifi Renani, A. M. Eustace, C. A. Myers, and C. W. Clary, "The Use of Synthetic IMU Signals in the Training of Deep Learning Models Significantly Improves the Accuracy of Joint Kinematic Predictions," *Sensors,* vol. 21, no. 17, p. 5876, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/17/5876.

[4] M. Błażkiewicz and A. Wit, "Artificial neural network simulation of lower limb joint angles in normal and impaired human gait," *Acta of bioengineering and biomechanics,* vol. 20, no. 3, pp. 43-49, 2018.

[5] M. Khant, D. Gouwanda, A. A. Gopalai, K. H. Lim, and C. C. Foong, "Estimation of Lower Extremity Muscle Activity in Gait Using the Wearable Inertial Measurement Units and Neural Network," *Sensors,* vol. 23, no. 1, p. 556, 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/1/556.

[6] U. Trinler, F. Leboeuf, K. Hollands, R. Jones, and R. Baker, "Estimation of muscle activation during different walking speeds with two mathematical approaches compared to surface EMG," *Gait & posture,* vol. 64, pp. 266-273, 2018.

[7] E. V. Zabre-Gonzalez, D. Amieva-Alvarado, and S. A. Beardsley, "Prediction of EMG activation profiles from gait kinematics and kinetics during multiple terrains," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2021: IEEE, pp. 6326-6329.

[8] N. F. Troje, "Decomposing biological motion: A framework for analysis and synthesis of human gait patterns," *Journal of Vision,* vol. 2, no. 5, pp. 2-2, 2002, doi: 10.1167/2.5.2.

[9] A. Daffertshofer, C. J. C. Lamoth, O. G. Meijer, and P. J. Beek, "PCA in studying coordination and variability: a tutorial," *Clinical Biomechanics,* vol. 19, no. 4, pp. 415-428, 2004/05/01/ 2004, doi: https://doi.org/10.1016/j.clinbiomech.2004.01.005.

[10] L. Van Der Maaten, E. Postma, and J. Van den Herik, "Dimensionality reduction: a comparative," *J Mach Learn Res,* vol. 10, no. 66-71, p. 13, 2009.

[11] J. Khodr and R. Younes, "Dimensionality reduction on hyperspectral images: A comparative review based on artificial datas," in *2011 4th International Congress on Image and Signal Processing*, 15-17 Oct. 2011 2011, vol. 4, pp. 1875-1883, doi: 10.1109/CISP.2011.6100531.

[12] M. Gavrilović and D. B. Popović, "A principal component analysis (PCA) based assessment of the gait performance," *Biomedical Engineering / Biomedizinische Technik,* vol. 66, no. 5, pp. 449-457, 2021, doi: doi:10.1515/bmt-2020-0307.

[13] J. W. Seo *et al.*, "Principal Characteristics of Affected and Unaffected Side Trunk Movement and Gait Event Parameters during Hemiplegic

Stroke Gait with IMU Sensor," (in eng), *Sensors (Basel),* vol. 20, no. 24, Dec 21 2020, doi: 10.3390/s20247338.
[14] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package)," *Neurocomputing,* vol. 307, pp. 72-77, 2018/09/13/ 2018, doi: https://doi.org/10.1016/j.neucom.2018.03.067.
[15] S. M. Moghadam, T. Yeung, and J. Choisne, "A comparison of machine learning models' accuracy in predicting lower-limb joints' kinematics, kinetics, and muscle forces from wearable sensors," *Scientific Reports,* vol. 13, no. 1, p. 5046, 2023/03/28 2023, doi: 10.1038/s41598-023-31906-z.
[16] S. Luqian and Z. Yuyuan, "Human Activity Recognition Using Time Series Pattern Recognition Model-Based on Tsfresh Features," in *2021 International Wireless Communications and Mobile Computing (IWCMC)*, 28 June-2 July 2021 2021, pp. 1035-1040, doi: 10.1109/IWCMC51323.2021.9498859.
[17] N. Kala, T. Bhatia, and N. Aggarwal, "Person Identification and Characterization from Gait Using Smartphone," in *2019 11th International Conference on Communication Systems & Networks (COMSNETS)*, 7-11 Jan. 2019 2019, pp. 492-495, doi: 10.1109/COMSNETS.2019.8711131.
[18] S. Arita, D. Nishiyama, T. Taniguchi, D. Fukui, M. Yamanaka, and H. Yamada, "Feature selection to classify lameness using a smartphone-based inertial measurement unit," *PLOS ONE,* vol. 16, no. 9, p. e0258067, 2021, doi: 10.1371/journal.pone.0258067.
[19] SENIAM. "SENIAM recommendations for determination of sensor location." http://seniam.org/ (accessed.
[20] J. J. Banks, W.-R. Chang, X. Xu, and C.-C. Chang, "Using horizontal heel displacement to identify heel strike instants in normal gait," *Gait & Posture,* vol. 42, no. 1, pp. 101-103, 2015/06/01/ 2015, doi: https://doi.org/10.1016/j.gaitpost.2015.03.015.
[21] I. T. J. Cadima and Jorge, "Principal component analysis: a review and recent developments," (in En), review-article 2016-4-13 2016, doi: doi:10.1098/rsta.2015.0202.
[22] J. Lever, M. Krzywinski, and N. Altman, "Principal component analysis," *Nature Methods,* vol. 14, no. 7, pp. 641-642, 2017/07/01 2017, doi: 10.1038/nmeth.4346.
[23] J. Korstanje, "The Random Forest," in *Advanced Forecasting with Python: With State-of-the-Art-Models Including LSTMs, Facebook's Prophet, and Amazon's DeepAR*. Berkeley, CA: Apress, 2021, pp. 179-191.
[24] "Principle component analysis of multivariate time series," in *Multivariate Time Series Analysis and Applications*, 2019, pp. 139-161.
[25] Y.-W. Kim, K.-L. Joa, H.-Y. Jeong, and S. Lee, "Wearable IMU-Based Human Activity Recognition Algorithm for Clinical Balance Assessment Using 1D-CNN and GRU Ensemble Model," *Sensors (Basel, Switzerland),* vol. 21, no. 22, p. 7628, 2021, doi: 10.3390/s21227628.
[26] J. M. Adams and K. Cerny, *Observational Gait Analysis: A Visual Guide*. Thorofare: Thorofare: SLACK, Incorporated, 2018.
[27] A. Schmitz, A. Silder, B. Heiderscheit, J. Mahoney, and D. G. Thelen, "Differences in lower-extremity muscular activation during walking between healthy older and young adults," *Journal of Electromyography and Kinesiology,* vol. 19, no. 6, pp. 1085-1091, 2009/12/01/ 2009, doi: https://doi.org/10.1016/j.jelekin.2008.10.008.