Clustering analysis of EMG cyclic patterns: A validation study across multiple locomotion pathologies

Clustering analysis of EMG cyclic patterns

A validation study across multiple locomotion pathologies

V. Agostini, S. Rosati, C. Castagneri, G. Balestra, M. Knaflitz
Department of Electronics and Telecommunications
Politecnico di Torino
Torino, Italy
valentina.agostini@polito.it

***Abstract*—Among other applications, electromyography (EMG) is used in the assessment of locomotion pathologies to quantitatively document abnormal muscle activation patterns during walking. However, EMG cyclic patterns are affected by high cycle-to-cycle variability. Previous research introduced a clustering approach (CIMAP) to recognize gait cycles with similar EMG onset/offset timings, reducing variability. To demonstrate the feasibility of the approach, the algorithm was validated on healthy subjects. The aim of this study is to extend the validation of the algorithm to multiple locomotion pathologies (both orthopedic and neurological). Gait data of a total of 50 subjects suffering from 5 different locomotion alterations were analyzed, considering 4 lower limb muscles. For each patient, datasets were built grouping EMG cycles with the same number of muscle activations. Then, hierarchical clustering was applied to each dataset and cycle-to-cycle variability was calculated for each cluster. Our results showed that CIMAP reduced the median variability below 5% of the gait cycle, for all the considered pathologies. Analyzing the number of clusters obtained we found that, in the great majority of cases, gait cycles cannot be bunched into a single group, but rather 2 or more clusters are necessary. As a consequence, the cluster representative elements, calculated by averaging cycles belonging to the same cluster, provide more trustworthy information for the clinician than indiscriminately averaging all cycles from a dataset.**

Keywords—clustering; gait analysis; EMG; human locomotion

I. INTRODUCTION

Electromyography (EMG) is used to assess musculoskeletal function during human cyclic movements (e.g. walking [1][2], running [3], cycling [4], swimming [5]), by recording muscle electrical activity in dynamic conditions. Among other applications, EMG is employed in the management of locomotion pathologies to quantitatively document abnormal muscle activation patterns during gait [6]. It may help clinicians to assess the effectiveness of therapeutic interventions and rehabilitation protocols in patients with orthopedic or neurological diseases compromising gait function [7][8], and in EMG biofeedback [9].

In order to obtain reliable gait measurements it is important to acquire long walking trials, to collect a high number of gait cycles [10]. For each gait cycle, muscle activation timing is obtained from the raw EMG signal, detecting when the signal is above background-noise level (ON) or below it (OFF) [11]. However, the high cycle-to-cycle variability of EMG onset/offset patterns makes it difficult to obtain synthetic and clear information for the clinician. This is one of the limitations to the spreading of EMG analysis in the clinical setting. More specifically, previous literature reported that 4-5 different muscle *activation modalities* are present in a subject's walk [12][13]. Each modality is characterized by a different number of activation intervals occurring within the gait cycle. Focusing on a single muscle, the muscle may be activated once, twice,…, or *n* times within the gait cycle, depending on the gait cycle considered along the walk. This cycle-to-cycle variability in EMG timing patterns was documented both in healthy and pathological subjects [14][15].

Moreover, a previous work highlighted that the muscle activation timing may significantly differ among cycles belonging to the same activation modality [16]. In other words, cycles in which the muscle activates the same number of times may present one or more of its activations in a different phase of the gait cycle, thus related to a different biomechanical task. Consequently, an erroneous interpretation of EMG signal may be produced if these cycles are averaged together.

For this reason, a clustering algorithm for identification of muscle activation patterns (CIMAP) was proposed and validated in young healthy individuals [16]. This allows for grouping cycles that share the same timing pattern, while keeping separate cycles with different onset-offset timing. However, the results of a validation on healthy individuals may not be extended to individuals with gait-related pathologies, since these present pathology-specific characteristics of gait variability.

The aim of this work is to validate the CIMAP algorithm in subjects with locomotion pathologies. To analyze a wide spectrum of gait-related pathologies, different kinds of orthopedic and neurological diseases were considered.

II. MATERIALS AND METHODS

A. Population and Experimental Protocol

Gait data of 50 pathological subjects were extracted from our data warehouse considering 5 different kinds of locomotion alterations. In particular we chose: 10 patients after total knee replacement (TKR), 10 patients after total hip replacement

(THR)[8], 10 elderly affected by normal pressure hydrocephalus (NPH)[17], 10 hemiplegic children Winters type I (W1) and 10 hemiplegic children hemiplegic type 2 (W2)[15]. TKR and THR are the consequence of orthopedic diseases, while NPH, W1 and W2 are neurological diseases.

All subjects walked barefoot, at self-selected speed, consecutively for 2-3 minutes, to allow the experimenter to collect a sufficient number of valid gait cycles for the analysis.

The wearable system STEP32 (Medical Technology, Italy) for clinical gait analysis was used to acquire gait data [18][10]. Foot-switches were placed under the foot-soles (size: 10 mm × 10 mm × 0.5 mm; activation force: 3 N), beneath the first and fifth metatarsal heads, and beneath the back portion of the heel.

Surface EMG probes were placed over tibialis anterior (TA), gastrocnemius lateralis (GL), rectus femoris (RF), and lateral hamstring (LH), bilaterally. These muscles were selected to study at least a pair of agonist-antagonist muscles acting at each joint of the lower limb. After skin preparation, probes were positioned over the muscles' belly (probe size: 27 mm × 19 mm × 7.5 mm, electrode: 4.0 mm diameter AgCl-disks, inter-electrode distance: 12 mm, gain: 1000). The sampling frequency was 2 kHz, and the EMG signal bandwidth ranged from 10 Hz to 450 Hz.

The experimental protocol conformed to the principles of the Helsinki declaration.

B. CIMAP Algorithm

In this study the CIMAP algorithm [16] is applied to EMG signal to group cycles showing similar activation patterns. CIMAP is a method based on agglomerative hierarchical clustering that consists of two steps: pre-processing and algorithm application.

In the first step, the EMG signal is segmented into separate gait cycles and processed to extract the onset and offset of each activation interval. Gait cycles with the same number of muscle activations m are grouped together to create a dataset (m^{th} activation modality).

In the second step, the dendrogram algorithm is applied to each dataset for obtaining the final clustering results. Clusters containing less than 5% of total gait cycles are considered non-

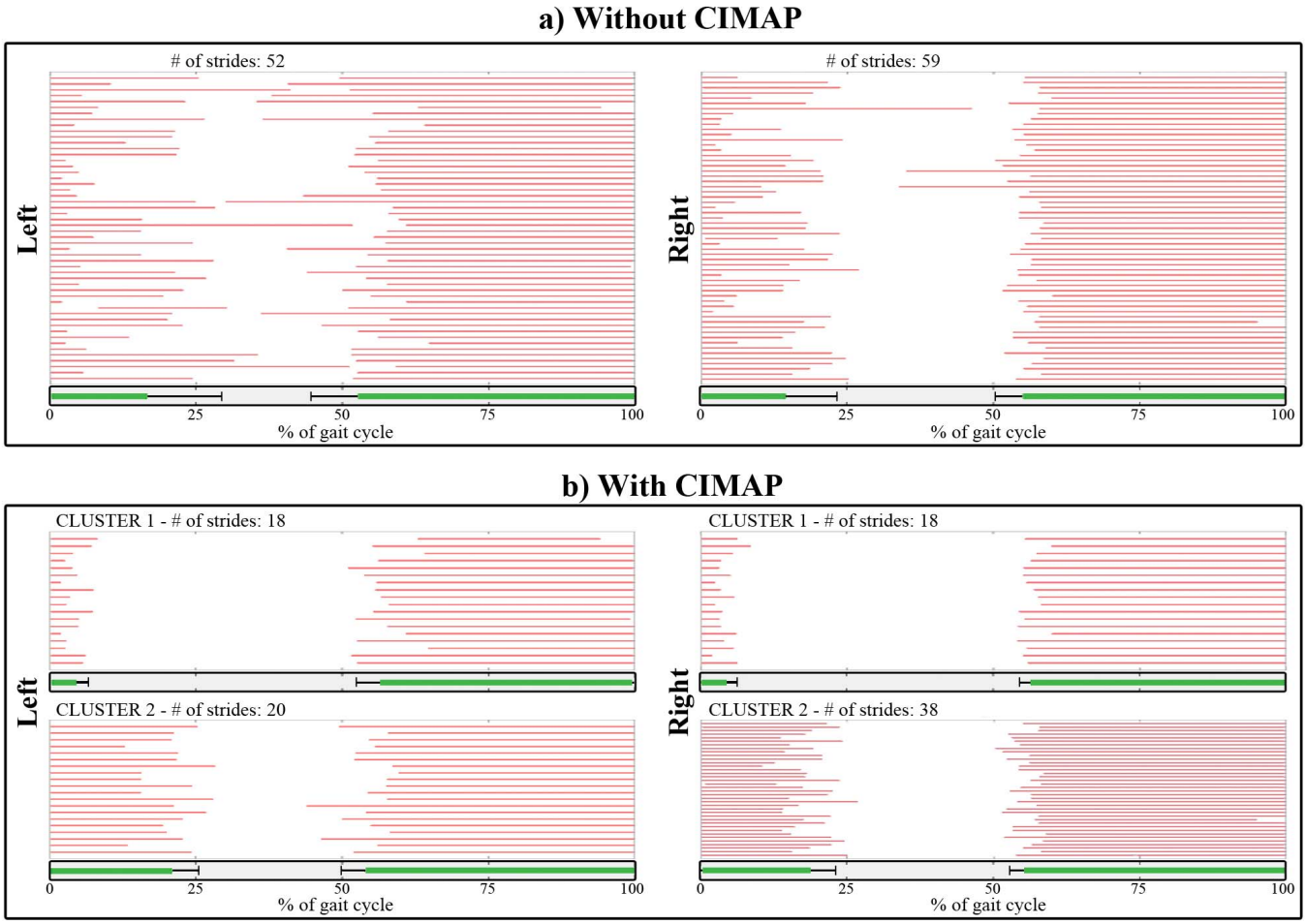

Fig. 1 Example of EMG activation intervals clustering (muscle: tibialis anterior). In each row, the red bar represents the EMG activation intervals within a gait cycle (RED = muscle active, WHITE = muscle inactive). The considered dataset contains 111 gait cycles (52 from the left and 59 from the right side) with 2 activation intervals (= 2nd activation modality). The green bar represents the average across gait cycles, the thin black line the corresponding standard deviation (SD). The dataset shows a high cycle-to-cyce variability (large SD) (a), that is reduced using CIMAP clustering algorithm (b).

significant and hence discarded. Moreover, for each valid cluster we calculate the centroid, averaging ON/OFF timings across the gait cycles belonging to a cluster. This centroid is used as representative element of the cluster. Differently from the previous version of the algorithm [16], in this study only datasets containing at least 35 gait cycles are processed by the clustering algorithm.

A more exhaustive description of the CIMAP algorithm can be found in [16].

C. Validation

In a previous work [16] the proposed method was validated on a population of subjects with no neurological or orthopedic pathologies that could influence their gait. In this study, the algorithm is applied to EMG signals acquired from subjects affected by 5 different locomotion pathologies. The results are compared with the situation in which no clustering is applied, in terms of cycle-to-cycle variability.

More specifically, for each obtained cluster we calculate the standard deviation (SD) of the *m* onset/offset timings, separately. Onsets at 0% of the gait cycle and offsets at 100% of the gait cycle are excluded since they show null SD. Then, we average the resulting SDs to obtain the cluster variability (*Cluster_Var*). The mean variability calculated for a specific subject is obtained by averaging *Cluster_Var* across all clusters, activation modalities and muscles.

Similarly, the variability without clustering (*NoCluster_Var*) was assessed considering all cycles sharing the same activation modality and calculating SDs of their *m* onset/offset timings. Also in this case, onsets at 0% and offsets

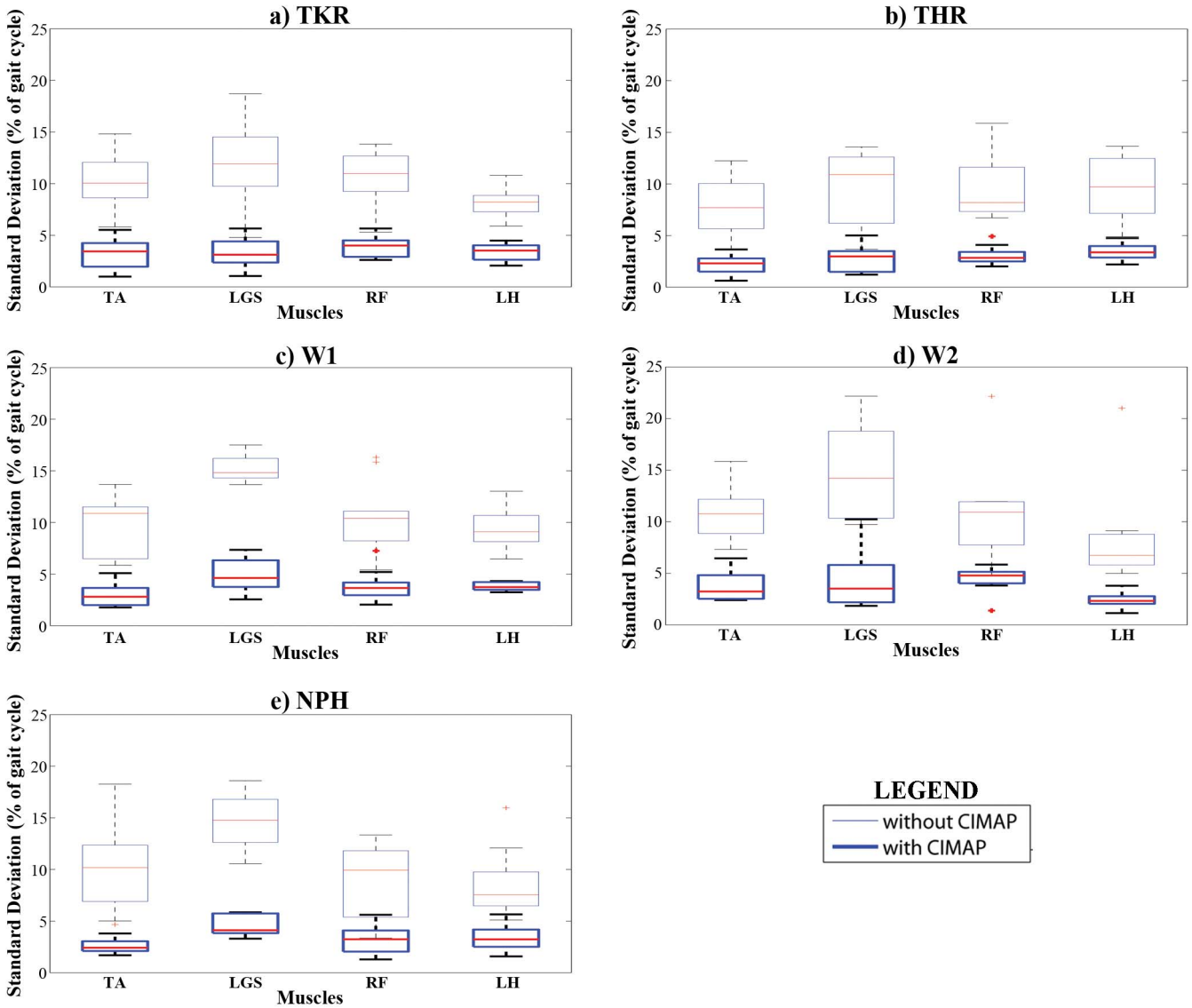

Fig. 2. Boxplots comparing the cycle-to-cycle variability before and after the application of the CIMAP algorithm, in 5 locomotion pathologies: (a) total knee replacement (TKR), (b) total hip replacement (THR), (c) hemiplegic children Winters type I (W1), (d) hemiplegic children Winters type II, (e) Normal Pressure Hydrocephalus (NPH). For each pathology the same 4 muscle were considered: Tibialis Anterior (TA), Gastrocnemius Lateralis (LGS), Rectus Femoris (RF), Lateral Hamstrings (LH).

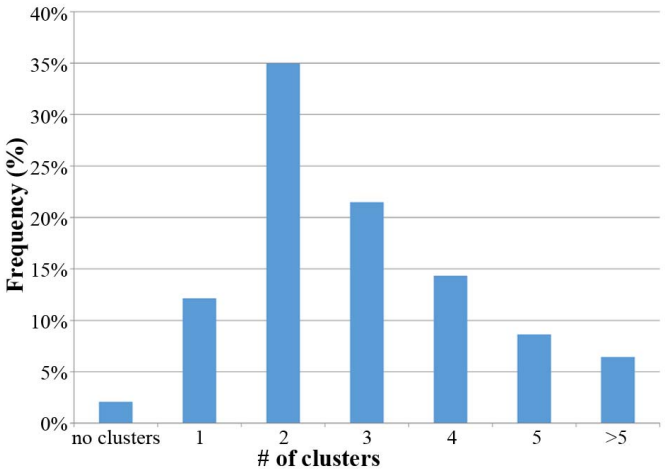

Fig. 3. Normalized histogram of the number of clusters obtained by CIMAP algorithm. The label "no clusters" indicates that no valid clusters were found.

at 100% of the gait cycle are discarded from the calculation. The mean variability for a specific subject is obtained by averaging *NoCluster_Var* across all activation modalities and muscles.

The variability estimated with and without clustering is represented by boxplots over the populations of subjects suffering from the 5 locomotion pathologies. Each pathology is considered separately, in order to assess its specific influence on the algorithm performance (if any).

Finally, we calculated the number of clusters obtained applying CIMAP algorithm to the whole sample (5 pathological populations pooled together). The normalized histogram of the number of clusters was built.

III. Results

For each subject, an average of 59±20 gait cycles (range: 35-151) was analyzed. The gait cycles were assigned to a variable number of datasets. Typically 1 to 4 datasets, each containing at least 35 cycles, were obtained for a specific subject's muscle. For two subjects (one THA and one W2) it was not possible to run the algorithm since there was no single dataset containing the minimum required number of cycles.

A total of 334 datasets were clustered by the CIMAP algorithm.

An example of the results obtained with the CIMAP algorithm applied to the 2-activation TA dataset of a representative subject is reported in Fig. 1. The figure compares the results obtained without clustering and with CIMAP (panels (a) and (b), respectively). In particular, the represented dataset initially shows a high cycle-to-cycle variability (Fig. 1a), but the variability is significantly reduced when separate clusters are considered after the CIMAP application (Fig. 1b).

The comparison of cycle-to-cycle variability with and without CIMAP is showed in Fig. 2, for each locomotion pathology. It is evident that, in all cases, the median variability is considerably reduced when using CIMAP, decreasing below 5% of the gait cycle.

Fig. 3 shows that, in the great majority of cases, the algorithm finds two or more clusters.

IV. Discussion

In this work the CIMAP algorithm was validated on 50 subjects suffering from different kinds of orthopedic and neurological diseases altering locomotion. For each of the considered pathology, the algorithm showed very good performances in grouping strides sharing similar EMG timing patterns. In particular, the cycle-to-cycle variability was considerably reduced after CIMAP application.

Analyzing the number of clusters obtained, it was evident that in the great majority of cases gait cycles cannot be bunched into a single group, but 2 or more clusters are necessary. This means that the representative element obtained averaging gait cycles from a dataset it is not really "representative" of the subject's gait, due to high cycle-to-cycle variability. On the contrary, the cluster representative elements (clusters centroids), calculated by averaging only gait cycles belonging to the same cluster, allows for obtaining more trustworthy information for the clinician.

On pathological subjects it is generally more difficult to perform long gait trials and, hence, to have available a high number of gait cycles. For this reason we had to lower the threshold defined in [16] (minimum number of gait cycles included in each dataset) from 50 to 35 gait cycles. Nevertheless, the algorithm performances, in terms of cycle-to-cycle variability reduction, are comparable to those obtained for healthy subjects [16]. However, using this threshold, it was not possible to run the algorithm for 2 out of 50 subjects. Although the algorithm threshold had to be lowered to only 35 cycles, the algorithm was able to find at least a valid cluster in 98% of cases.

V. Conclusions

The CIMAP algorithm for identification of EMG activation patterns was validated considering different kinds of locomotion pathologies (both orthopedic and neurological). This algorithm reduces the cycle-to-cycle variability providing more representative EMG patterns and, hence, giving clearer and more complete information to the clinician.

This study focused on the cyclic task of walking. However, the proposed approach can be easily extended to others cyclic movements.

References

[1] D. Winter and H. Yack, "EMG profiles during normal human walking: stride-to-stride and inter-subject variability," Electroencephalography and clinical neurophysiology, 1987.

[2] F. Sylos-Labini, V. La Scaleia, A. d'Avella, I. Pisotta, F. Tamburella, G. Scivoletto, M. Molinari, S. Wang, L. Wang, E. van Asseldonk, H. van der Kooij, T. Hoellinger, G. Cheron, F. Thorsteinsson, M. Ilzkovitz, J. Gancet, R. Hauffe, F. Zanov, F. Lacquaniti, and Y. P. Ivanenko, "EMG patterns during assisted walking in the exoskeleton," Frontiers in Human Neuroscience, vol. 8, no. June, p. 423, 2014.

[3] E. S. Chumanov, C. M. Wille, M. P. Michalski, and B. C. Heiderscheit, “Changes in muscle activation patterns when running step rate is increased.,” Gait & posture, vol. 36, no. 2, pp. 231–5, Jun. 2012.

[4] S. K. Sidhu, B. W. Hoffman, A. G. Cresswell, and T. J. Carroll, “Corticospinal contributions to lower limb muscle activity during cycling in humans.,” Journal of neurophysiology, vol. 107, no. 1, pp. 306–14, Jan. 2012.

[5] J. Martens, D. Daly, K. Deschamps, F. Staes, and R. J. Fernandes, “Inter-individual variability and pattern recognition of surface electromyography in front crawl swimming.,” Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology, vol. 31, pp. 14–21, Sep. 2016.

[6] J. Perry, Gait Analysis: Normal and Pathological Function. Thorofare, New Jersey: SLACK Incorporated, 1992.

[7] M. Benedetti, V. Agostini, M. Knaflitz, V. Gasparroni, M. Boschi, and R. Piperno, “Self-reported gait unsteadiness in mildly impaired neurological patients: an objective assessment through statistical gait analysis,” Journal of NeuroEngineering and Rehabilitation, vol. 9, no. 1. p. 64, 2012.

[8] V. Agostini, D. Ganio, K. Facchin, L. Cane, S. Moreira Carneiro, and M. Knaflitz, “Gait parameters and muscle activation patterns at 3, 6 and 12 months after total hip arthroplasty.,” The Journal of arthroplasty, vol. 29, no. 6, pp. 1265–1272, Jun. 2014.

[9] H. Woodford and C. Price, “EMG biofeedback for the recovery of motor function after stroke,” Cochrane database of systematic reviews (Online), no. 2, p. CD004585, 2007.

[10] V. Agostini, F. Lo Fermo, G. Massazza, and M. Knaflitz, “Does texting while walking really affect gait in young adults?,” Journal of NeuroEngineering and Rehabilitation, vol. 12, no. 1, p. 86, Sep. 2015.

[11] P. Bonato, T. D’Alessio, and M. Knaflitz, “A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait,” IEEE Transactions on Biomedical Engineering, vol. 45, pp. 287–299, 1998.

[12] V. Agostini, A. Nascimbeni, A. Gaffuri, P. Imazio, M. G. Benedetti, and M. Knaflitz, “Normative EMG activation patterns of school-age children during gait,” Gait and Posture, vol. 32, no. 3, pp. 285–289, 2010.

[13] F. Di Nardo, A. Mengarelli, L. Burattini, E. Maranesi, V. Agostini, A. Nascimbeni, M. Knaflitz, and S. Fioretti, “Normative EMG patterns of ankle muscle co-contractions in school-age children during gait,” Gait & Posture, vol. 46, pp. 161–166, May 2016.

[14] F. Di Nardo, E. Maranesi, A. Mengarelli, G. Ghetti, L. Burattini, and S. Fioretti, “Assessment of the variability of vastii myoelectric activity in young healthy females during walking: A statistical gait analysis,” Journal of Electromyography and Kinesiology, vol. 25, no. 5, pp. 800–807, 2015.

[15] V. Agostini, A. Nascimbeni, A. Gaffuri, and M. Knaflitz, “Multiple gait patterns within the same Winters class in children with hemiplegic cerebral palsy,” Clinical Biomechanics, vol. 30, no. 9, pp. 908–914, 2015.

[16] S. Rosati, V. Agostini, M. Knaflitz, and G. Balestra, “Muscle activation patterns during gait: A hierarchical clustering analysis,” Biomedical Signal Processing and Control, vol. 31, pp. 463–469, 2017.

[17] V. Agostini, M. Lanotte, M. Carlone, M. Campagnoli, I. Azzolin, R. Scarafia, G. Massazza, and M. Knaflitz, “Instrumented gait analysis for an objective pre/post assessment of tap test in normal pressure hydrocephalus,” Archives of Physical Medicine and Rehabilitation, vol. 96, no. 7, pp. 1235–1241, 2015.

[18] V. Agostini, M. Knaflitz, L. Antenucci, G. Lisco, L. Gastaldi, and S. Tadano, “Wearable sensors for gait analysis,” in 2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings, 2015, pp. 146–150. DOI:10.1109/MeMeA.2015.7145189.