Article

A Supervised Classification of Children with Fragile X Syndrome and Controls Based on Kinematic and sEMG Parameters

Weronika Joanna Piatkowska [1], Fabiola Spolaor [1], Marco Romanato [1], Roberta Polli [2,3], Alessandra Huang [1], Alessandra Murgia [2,3,†] and Zimi Sawacha [1,4,*,†]

1   Department of Information Engineering, University of Padova, 35131 Padova, Italy; weronikajoanna.piatkowska@studenti.unipd.it (W.J.P.); fabiola.spolaor@unipd.it (F.S.); romanato@dei.unipd.it (M.R.); alessandra.huang@studenti.unipd.it (A.H.)
2   Department of Women's and Children's Health, University of Padova, 35128 Padova, Italy; roberta.polli@unipd.it (R.P.); alessandra.murgia@unipd.it (A.M.)
3   Istituto di Ricerca Pediatrica CDS, 35127 Padova, Italy
4   Department of Medicine, DIMED, University of Padova, 35128 Padova, Italy
*   Correspondence: zimi.sawacha@dei.unipd.it
†   These authors contributed equally to this work.

Abstract: Fragile X syndrome (FXS) is caused by pathologic expansions of the CGG repeat polymorphic region of the FMR1 gene. There are two main categories of FMR1 mutations, "premutation" and "full mutation", that are associated with different clinical phenotypes, and somatic mosaicism can represent a strong FXS phenotype modulator. FXS is the leading cause of inherited intellectual disability and autism, and it is characterized by musculoskeletal manifestations such as flexible flat feet, joint laxity and hypotonia. The former have been associated with altered joint kinematics and muscle activity during gait. The aim of this study was to use gait analysis parameters to classify FXS children from healthy controls and, within FXS children with full mutation, to classify children with mosaicism. Seven supervised machine learning algorithms were applied to a dataset of joint kinematics and surface electromyographic signals collected on twenty FXS children and sixteen controls. Results showed that the k-NN algorithm outperformed in terms of accuracy (100%) in classifying FXS children from controls, while CN2 rule induction obtained the best accuracy (97%) in classifying FXS children with mosaicism. The proposed pipeline might be used for developing assisted decision-making systems aiming at identifying and treating the musculoskeletal alterations associated with FXS.

Keywords: supervised learning; fragile X syndrome; gait analysis; surface electromyography; kinematics

1. Introduction

Fragile X syndrome (FXS) is the leading cause of inherited intellectual disability and autism (frequency estimated at 1/4000–1/7000) [1] caused by pathologic expansions of the CGG repeat polymorphic region of the FMR1 gene. The number of the CGG repeats determines two main mutational categories with different clinical phenotypes—"premutation" and "full mutation". In the premutation (55–200 CGGs) different phenotypes such as fragile X-associated primary ovarian insufficiency, fragile X-associated tremor ataxia syndrome and fragile X associated neuropsychiatric disorders [2,3] are caused by the cell toxicity caused by an excessive production of the FMR1 gene. In the full mutation (>200 CGGs) FXS phenotype with symptoms such as intellectual disability, behavioural deficits and global development delays [4] is caused by gene silencing and the loss of protein function. In this context, somatic mosaicism can be a strong FXS phenotype modulator [4]. There are two types of mosaicism—"size mosaicism", when both premutation cells and full mutation

cells are present, leading to the occurrence of problems connected with both mutational categories; and "methylation mosaicism", when full mutation expansions are not completely methylated [5].

In FXS children, musculoskeletal manifestations such as flexible flat feet, joint laxity and hypotonia [6] may lead to non-physiological gait patterns [6,7]. Gait analysis in FXS children has documented significant alterations such as a characteristic pattern of excessively flexed hip and ankle joints with reduced flexion at the knee, suggesting an overall immature motor control [7,8]. In terms of surface electromyography, in one study an altered muscle activity was detected, and an asymmetric activation pattern on gastrocnemius lateralis and biceps femoris suggested the presence of compensatory mechanisms actuated by these children to cope with the overall joint laxity and flexible flat feet [7].

In the recent literature, gait patterns associated with various pathologies (i.e., cerebral palsy patients, Parkinsonians, post-stroke patients) have been investigated in the context of gait classification to assist with clinical decision making [9,10]. Supervised machine learning algorithms operate on a training set of input data (e.g., gait parameters) and known labels of the data (e.g., clinical diagnosis) to train a predictive model to be able to generate accurate predictions for new unlabelled data. Kamruzzaman et al. in 2007 applied the support vector machine algorithm to distinguish children with cerebral palsy from healthy controls based on kinetic gait parameters, and achieved 97% accuracy [11]. Zhang et al. in 2019 used kinematic features as an input for seven different classifiers, and reported the best accuracy for an artificial neural network [12]. Mannini et al. in 2016 applied the support vector machine algorithm with gait characteristics obtained from inertial measurement units as an input to classify post-stroke patients, Huntington's disease patients and healthy elderly individuals with an accuracy of 90.5% [13]. Lee et al. in 2012 used gait characteristics and wavelet-based features as an input for support vector machine to classify Parkinsonian patients and healthy controls with 77.3% accuracy. According to the latest review by Figueiredo et al., the application of machine learning to the recognition of gait patterns is likely to provide a potentially realistic diagnosis [9].

However, in the context of FXS, to the best of the authors' knowledge, automatic gait classification has not previously been performed. The aim of the present study was to use gait analysis parameters to classify FXS children from healthy controls and, within FXS children with full mutation, to classify children with mosaicism. Toward this aim, seven supervised machine learning algorithms were applied to a dataset of joint kinematics and surface electromyographic signals collected on twenty FXS children and sixteen controls. Understanding the key gait analysis variables which discriminate FXS children from controls with typical neurodevelopment and, within FXS, children carrying somatic mosaicism, has the potential to promote the identification and treatment of the musculoskeletal conditions related to this specific genetic condition.

2. Materials and Methods

2.1. Participants

Appropriate informed consent was obtained for each subject enrolled in the study for participation, scientific use of the data and publication (Local Ethics Committee, Università Azienda Ospedaliera di Padova, trial n° 46039, registration date 29 July 2019).

Twenty male children with FXS phenotype (mean ($\pm$SD) age and BMI, respectively, of 10.00 ($\pm$3.74) years and 18.57 ($\pm$3.46) kg/m^2) were evaluated in a routine clinical setting at the Department of Women's and Children's Health, University of Padua; thirteen of the FXS children carried a full mutation of the FMR1 gene (FXS-FM; age of 10.69 ($\pm$3.67) years, BMI of 18.82 ($\pm$4.09) kg/m^2), five of them carried a methylation mosaicism (FXS-M; age of 9.00 ($\pm$3.35) years, BMI of 18.70 ($\pm$2.34) kg/m^2), two carried a premutation (FXS-P) and were affected with a phenotype compatible with FXS. Sixteen controls with typical neurodevelopment (CS; mean ($\pm$SD) age of 10.22 ($\pm$3.19) years, BMI of 22.84 ($\pm$3.93) kg/m^2) were evaluated at the BioMovLab (University of Padua, Padova, Italy).

Subjects' numerosity was defined based on power analysis [14] carried out on a previously published dataset [7] considering the value of the envelope peak as a variable; a number of 11 subjects was found to be sufficient for our analysis, and our sample consisted of 18 FXS children and 16 controls.

FXS subjects were enrolled according to the following inclusion criteria:

- Molecularly documented full mutation of the FMR1 gene with expansions of more than 200 CGG repeats and methylation of the promoter and repeated sequence; possible size and/or methylation mosaicism.
- Ability to walk independently.
- Absence of documented orthopaedic comorbidities affecting the lower limbs within 12 months from the beginning of the study.
- Absence of documented neurological disorders.

BMI- and age-matched subjects with typical neurodevelopment were enrolled as controls for the study according to the following inclusion criteria:

- Ability to walk independently.
- Absence of documented lower-limb injures within 12 months of the beginning of the study.
- Absence of documented neurological disorders.

2.2. Molecular Analysis

Molecular analysis was performed as previously reported [7]. Genomic DNA (gDNA) was extracted from peripheral blood leukocytes (PBLs) and saliva on an automated Maxwell® 16 Blood DNA Purification System (Promega, Milan, Italy) and quantified by spectrophotometer with NanoDrop™ (ThermoFisher Scientific, Waltham, MA, USA). To identify the full range of FMR1 CGG repeat expansions, genomic DNA (40–60 ng) was amplified with an Amplidex FMR1 PCR kit (Asuragen, Austin, TX, USA) as previously described [15] and according to the manufacturer's recommended protocol. All amplicons were analysed by capillary electrophoresis (CE) on a 3130*xl* Genetic Analyzer (Applied Biosystems, ThermoFisher Scientific, Waltham, Massachusetts, USA). Fragment length was derived from the size of the PCR products calibrated to a ROX 1000 Size Ladder (Asuragen, Austin, TX, USA) with the use of the GeneMapper® Software v 4.0 (ThermoFisher Scientific, Waltham, MA, USA). Methylation analysis was carried out with use of the AmplideX FMR1 mPCR kit (Asuragen, Austin, TX, USA), as described [16]. Methylation percentage was calculated as a ratio of peak heights between digested (HEX) and undigested samples (FAM), normalized to the CGG control amplicon peak height with the GeneMapper® Software v 4.0. Alleles are reported as unmethylated (<20%), partially methylated (20–80%) and fully methylated (>80%).

2.3. Instrumental Assessment

Kinematics and sEMG data were simultaneously acquired through four synchronized cameras (GoPro Hero3, 30 fps) and an sEMG system (FreeEmg, BTS, 1000 Hz) that collected the activity of tibialis anterior (TA), gastrocnemius lateralis (GL), rectus femoris (RF) and biceps femoris (BF). sEMG sensors were positioned according to [17,18] after appropriately cleaning and preparing the skin. Sensors were 24 mm in diameter and positioned 1 cm apart. Each subject performed several gait trials at a self-selected speed; at least three trials per subject were processed.

The kinematic data were acquired as described in [7] without applying any markers on subjects' skin. Video sequences were processed to extract the three-dimensional anatomical landmark coordinates using the software "Track on Field" (BBSoF s.r.l.), which implements the optical flow popular Kanade–Lucas–Tomasi tracking algorithm [19] in the version proposed by Sawacha et al. (2014) [20,21]. The algorithm was modified and validated as described in [7] to enable tracking with the markers automatically added in anatomical landmarks identified by the user (Figure 1). Data were processed in Matlab

(MathWorks v.2019a) and sagittal plane kinematics as well as spatiotemporal parameters were extracted according to a simplified version of [19] as in [20]. Minimal and maximal values of joint angles (hip, knee and ankle) and joints' range of motion were calculated using self-developed Matlab codes.
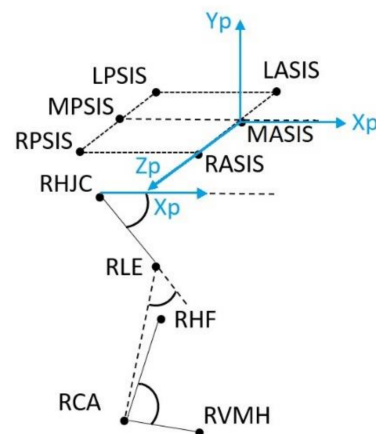

Figure 1. The definition of lower limb embedded frames and sagittal plane angles. The pelvis reference system is represented in blue: the origin is the midpoint between LASIS and RASIS (MASIS), the z axis is the normalized vector oriented as the line passing through the LASIS and RASIS with its positive direction from left to right; the x axis lies in the plane defined by the RASIS, LASIS and the midpoint between the RPSIS and LPSIS (MPSIS) with its positive direction forwards; the y axis is orthogonal to the xz plane and its positive direction is proximal. The hip flexion angle is defined as the angle between pelvis and femur (femur is the line connecting the hip joint centre (HJC) and the lateral epicondyle (LE)); the knee flexion angle is defined as the angle between femur and shank (shank is the line connecting the LE and the calcaneus (CA)); the ankle flexion angle is defined as the angle between shank and foot (foot is the line connecting the CA and the fifth metatarsal head). Abbreviations: right and left posterior superior iliac spines (RPSIS, LPSIS); right and left anterior superior iliac spines (RASIS, LASIS); right lateral epicondyle (RLE); right head of fibula (RHF); right calcaneus (RCA); right fifth metatarsal head (RVMH); right HJC (RHJC).

In terms of sEMG analysis the following parameters were computed: duration of contraction, onset and offset activation timing, peak of the envelope and occurrence in the gait cycle, muscle co-contractions number and occurrence in the gait cycle [21–25].

The envelope was computed by low-pass filtering the signals with a fourth-order Butterworth filter and a cut-off frequency of 5 Hz. The peak of each muscle sEMG activity was determined as a percentage of medium value in the cycle (PoE%) as well as its occurrence with respect to the gait cycle (PPoE%) [21,23].

To obtain the activation timing, signals were filtered by filter removing heartbeat, notch filter for 50 Hz, band-pass filtered with a double fifth-order Butterworth filter and full wave rectified. The cut-off frequencies varied from 5 to 15 Hz for the high-pass filter, and from 450 to 495 Hz for the low-pass filter. A double-threshold statistical detector proposed by Bonato [22] was applied. Background noise was estimated for each signal based on the interval of the subject's static standing. Only activation intervals longer than 30 ms were accepted [22]. The values of duration of muscle activity and the onset and offset of muscle activity intervals were detected for each gait cycle.

In order to calculate the co-contraction ratio (CCR) for knee and ankle joints, Equation (1) was implemented [25]. Antagonist muscles for the knee joint were BF and GL, while RF was the agonist. For the ankle joint the agonist muscle was TA, while GL was the antagonist. An index of −1 indicates a sole contribution of the agonist excitations, 1 indicates a contribution derived just from antagonist excitations and 0 indicates a balanced contribution of both agonist and antagonist excitations (i.e., the maximum value of co-contraction).

$EMG(t)_{antagonist}$ and $EMG(t)_{agonist}$ were respectively representative of the average of all the agonist and antagonist muscle excitations in the same instant of time.

$$
\begin{cases}
CCR(t) = \dfrac{EMG(t)_{antagonist}}{EMG(t)_{agonist}} - 1, & if\ EMG(t)_{antagonist} < EMG(t)_{agonist} \\
CCR(t) = 1 - \dfrac{EMG(t)_{antagonist}}{EMG(t)_{agonist}}, & if\ EMG(t)_{antagonist} > EMG(t)_{agonist}
\end{cases}
\tag{1}
$$

To extract a value that could be representative for the co-contractions (CCs) in our cluster analysis, we defined a CC to occur at $CCR(t)$ equal to 0. We then computed the number of CCs and extracted their occurrence within the gait cycle.

The extracted parameters are listed in the following.

sEMG parameters:

- Duration of muscle contraction;
- Muscle activation onset and offset;
- Value of envelope peak;
- Envelope peak occurrence within the gait cycle.

Co-contraction parameters:

- Number of co-contractions within the gait cycle;
- Co-contraction occurrence within the gait cycle.

Kinematic parameters:

- Minimal values of joints' articular angles (hip, knee and ankle);
- Maximal values of joints' articular angles (hip, knee and ankle);
- Range of motion of joints' articular angles (hip, knee and ankle).

2.4. Classification Procedure

The overview of the procedure for the gait pattern classification is presented in Figure 2. The input data vectors for classification algorithms were formed from the elaborated parameters as follows:

- S1: sEMG parameters and co-contraction parameters;
- S2: kinematic parameters;
- S3: sEMG parameters, co-contraction parameters and kinematic parameters.

Seven supervised learning algorithms [26] were applied as a gait classification system: decision tree, random forest, CN2 rule induction, support vector machine (SVM), k-nearest neighbours (k-NN), neural network and naïve Bayes. All the algorithms were implemented in Orange Mining Toolbox.

Two types of analysis were performed: first the classification between FXS subjects and controls was carried out using three different sets of vectors (S1, S2 and S3); then, the classification between different mutation categories of the FMR1 gene (FXS-FM vs. FXS-M) using two different sets of input vectors (S1 and S2) was performed.

For each analysis the training set consisted of a randomly selected set representing 66% of the gait trials and the test set consisted of the remaining 33% of gait trials. For the training set a 10-fold cross-validation was performed.

The extracted parameters were statistically analysed by means of Kruskal–Wallis non-parametric test ($p < 0.05$) in SPSS (v24, IBM Statistical Toolbox) due to the non-normal distribution of the data according to the Shapiro–Wilk test.
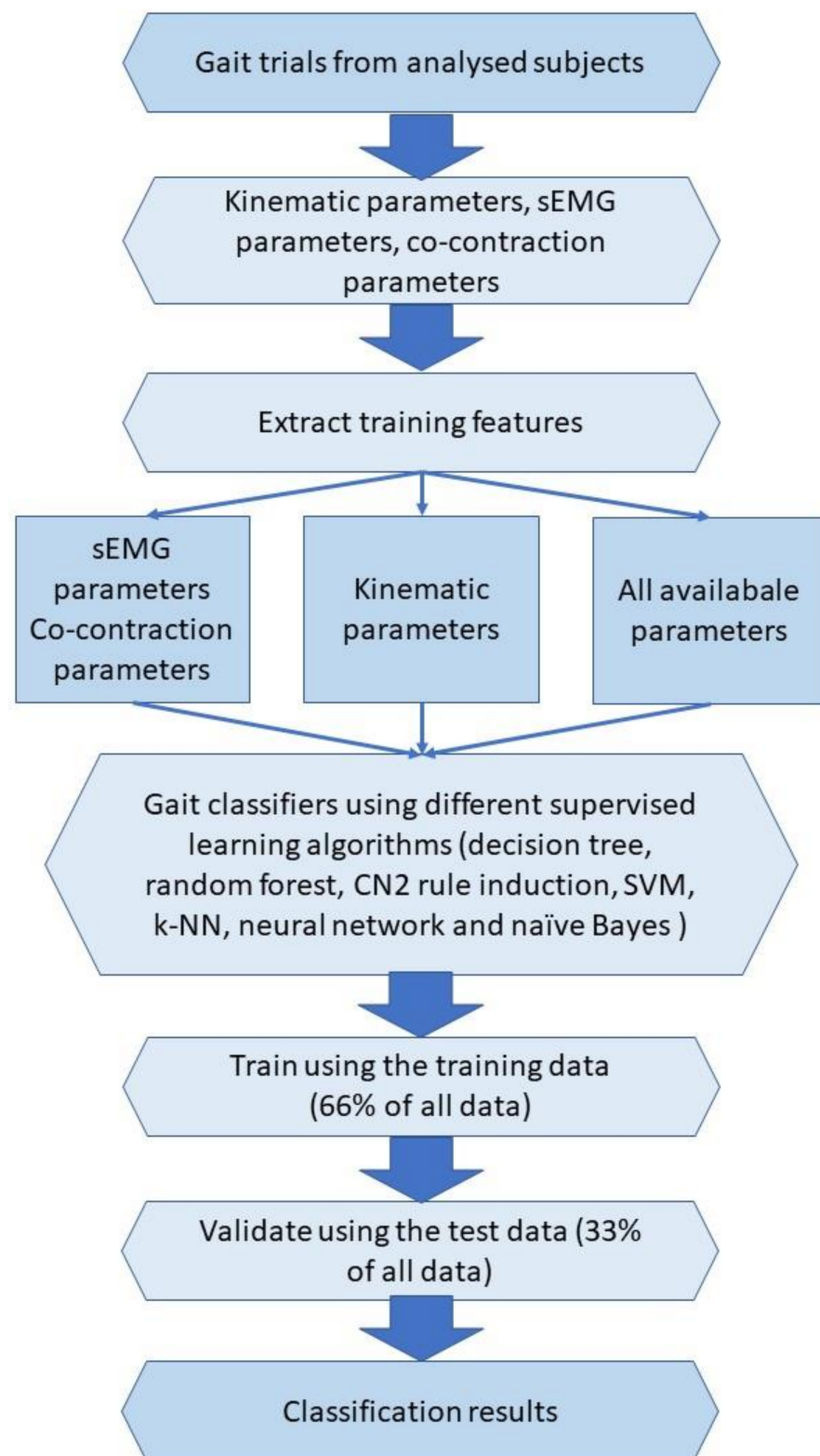
Figure 2. Diagram with the overview of the gait classification using seven supervised learning algorithms.

3. Results

3.1. FXS vs. CS

The classification accuracy for classification between FXS and CS in all the analysed supervised algorithms is summarized in Table 1.

Table 1. Accuracy of the seven supervised learning algorithms for the classification between FXS and CS.

Input Vector/ Classification Algorithm	Classification Accuracy					
	S1		S2		S3	
	Training [a]	Test [b]	Training [a]	Test [b]	Training [a]	Test [b]
Random Forest	98.3%	98.3%	99.1%	98.3%	99.2%	96.7%
CN2 Rule Induction	96.2%	94.2%	97.4%	98.3%	95.8%	92.7%
Neural Network	95.7%	94.2%	99.1%	96.6%	97.9%	95.9%
SVM	97.9%	95.0%	100.0%	98.3%	97.9%	98.4%
Decision Tree	94.9%	97.5%	96.5%	98.3%	93.8%	90.2%
k-NN	95.3%	95.0%	100.0%	100.0%	96.3%	95.9%
Naïve Bayes	88.9%	90.8%	99.1%	96.6%	77.5%	76.4%

S1: sEMG parameters and co-contraction parameters. S2: kinematic parameters. S3: sEMG parameters, co-contraction parameters and kinematic parameters. [a] Classification accuracy on the training dataset. [b] Classification accuracy on the test dataset.

Overall, all the algorithms were generally able to classify gait characteristics from different input vectors. The average predictive accuracy for new data was 95.0% for S1, 98.1% for S2 and 92.3% for S3. The best accuracy for S1 was achieved using random forest (98.3%); for S2 using k-NN (100%) and for S3 using SVM (98.4%). The worst accuracy for all analysed vector sets was obtained using naïve Bayes.

In terms of input parameters, the best classification was obtained using S2 (kinematic parameters). This subset of parameters was then further investigated in order to verify the presence of statistically significant differences between the FXS and CS populations of subjects, and statistically significant differences ($p < 0.05$) were found in maximum hip, knee and ankle angles; minimum hip and knee angles; and hip and knee range of motion (see Figure 3).
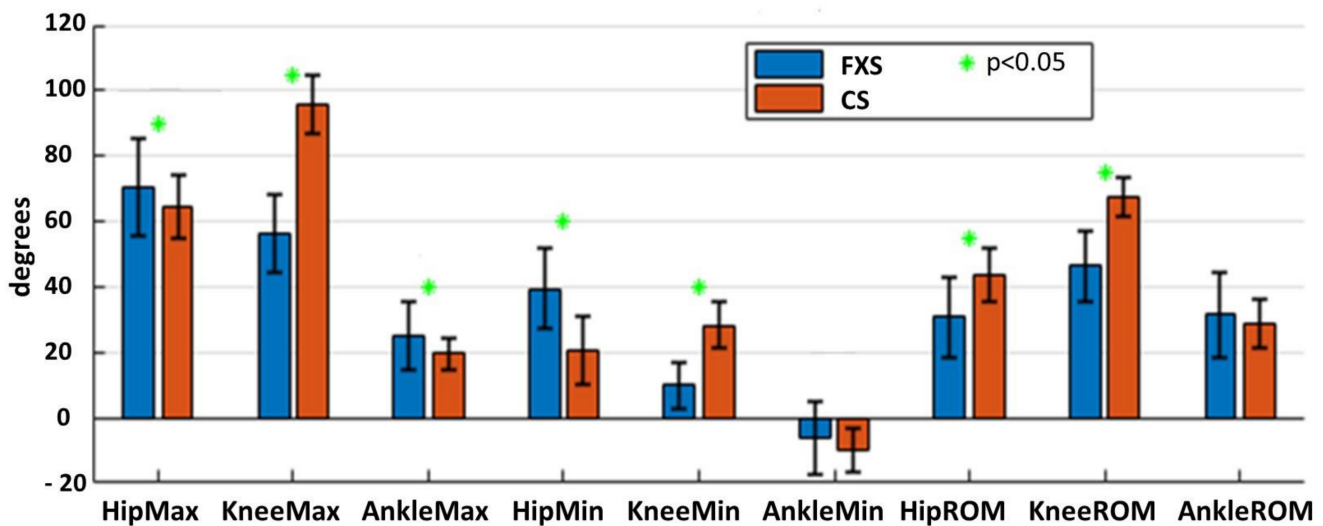

Figure 3. Kinematic parameters used as input for the classification algorithms (S2) in FXS and CS subjects. * Statistically significant ($p < 0.05$).

3.2. FXS-FM vs. FXS-M

The accuracy for the classification of FXS-M within the FXS-FM group is summarized in Table 2.

Table 2. Accuracy of the seven supervised learning algorithms for the classification of FXS-M within the FXS-FM group.

Input Vector/ Classification Algorithm	Classification Accuracy			
	S1		S2	
	Training [a]	Test [b]	Training [a]	Test [b]
Random Forest	96.9%	90.9%	78.5%	78.8%
CN2 Rule Induction	96.9%	97.0%	63.1%	72.7%
Neural Network	98.5%	90.9%	83.1%	84.8%
SVM	95.4%	87.9%	72.3%	69.7%
Decision Tree	96.9%	93.9%	78.5%	63.6%
k-NN	94.6%	93.9%	80.0%	84.8%
Naïve Bayes	73.8%	80.3%	61.5%	69.7%

S1: sEMG parameters and co-contraction parameters. S2: kinematic parameters. [a] Classification accuracy on the training dataset. [b] Classification accuracy on the test dataset.

Overall, all the algorithms were generally able to classify gait characteristics from different input vectors. The average predictive accuracy for new data was 90.7% for S1 and 74.9% for S2. The best accuracy for S1 was achieved using CN2 rule induction (97.0%) while for S2 it was obtained using k-NN and neural network (84.8% for both algorithms). The worst accuracy for both vector sets was obtained using naïve Bayes.

In terms of input parameters, the best classification was obtained using S1 (sEMG and co-contraction parameters). This subset of parameters was then further investigated in order to verify the presence of statistically significant differences between FXS-M and FXS-FM. Statistically significant differences were observed in the position of the envelope peak in the left TA, GL, Bf and both the right TA and RF; the value of the envelope peak in the left TA and the right TA, GL and BF; the muscle duration in GL; the muscle activity onset and offset in both the left RF and BF; and the number of co-contractions in the left ankle joint and the right knee joint (see Figure 4).
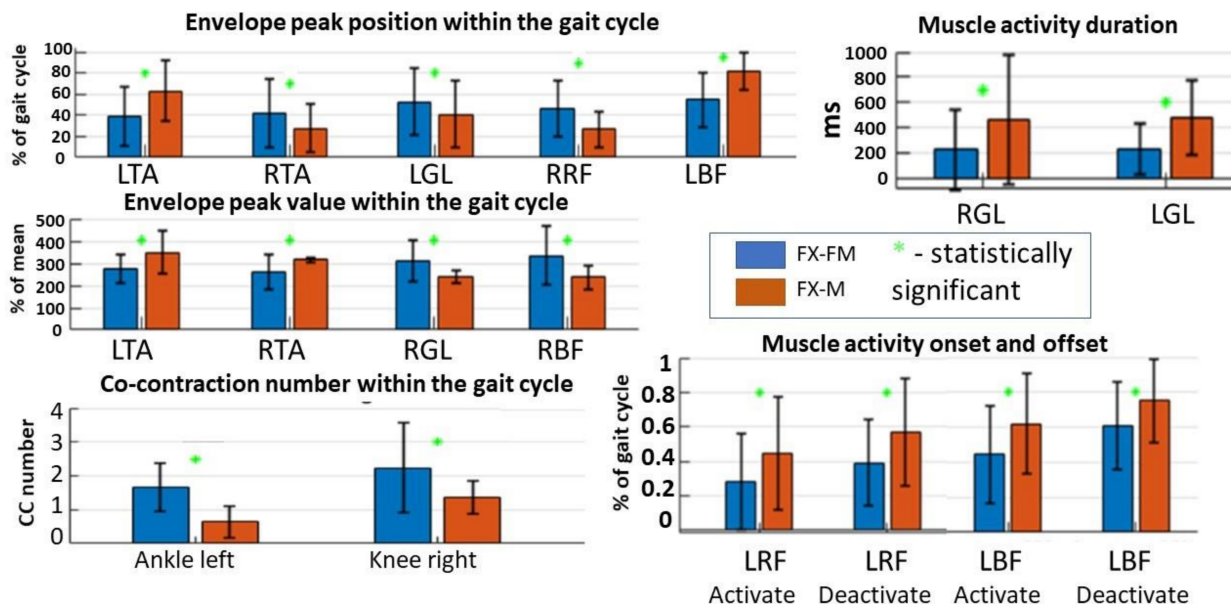

Figure 4. sEMG and co-contraction parameters in which statistically significant difference ($p < 0.05$) was found between FXS-FM and FXS-M.

4. Discussion

The current state of the art is focusing on the development of automatic techniques to discriminate pathological gait patterns from healthy ones, in order to track disease progression and prescribe target therapies [4]. In this context, machine learning algorithms

have shown the ability on one side to identify atypical gait patterns and on the other to manage multidimensional data with the possibility of increasing dataset size without the need to change the methodological approach [9,27]. In the present contribution the standard machine learning algorithms were applied to an under-investigated pathology (FXS) from a gait-analysis perspective. In the majority of the state-of-the-art works spatiotemporal gait features have been used for gait pattern classification, accompanied by 3D gait joint angles and kinetics features [9]. Previous studies applied supervised learning algorithms driven by kinematic data acquired during gait to differentiate subjects with chronic neck pain and healthy ones [28], predict the severity of Parkinson's disease [29], categorize gait patterns in CP subjects [12,30] and link toe-walking gait patterns to their clinical cause [31]. In the present contribution, in contrast to previous works adopting only kinematic or only sEMG features, both types of features were combined in different ways in order to fully explore their classification potential. Furthermore, to the best of our knowledge, co-contraction indexes were included for the first time in the attempt to classify gait patterns.

With respect to the study aim, our results support the use of gait analysis parameters to classify FXS children and, within FXS children with full mutation, to classify children with mosaicism. In this context, the overall highest accuracy for the classification between FXS and CS subjects was found using the k-NN algorithm (100%), while for the classification between FXS-M and FXS-FM, CN2 rule induction yielded the best result (Table 2).

The classification accuracy for the other algorithms (76.4–98.3% for FXS vs. CS and 69.7–93.9% for FXS-FM vs. FXS-M) was in the same range reported by previous literature in the domain of gait classification [26,32]. The differences in terms of performance among the different algorithms were generally small (up to 6.5% for all the analysed algorithms except for naïve Bayes and up to 1.7% for random forest, k-NN, SVM and decision tree in FXS vs. CS classification).

The obtained results suggest that random forest, k-NN, SVM and decision tree are good candidates for future use in the development of a classification system for discriminating FXS children from CS based on their gait patterns. Meanwhile, when considering FXS-M, the CN2 rule induction was the best candidate for developing a classification system aiming to distinguish FXS-M within a group of FXS-FM individuals.

The present contribution not only revealed the most useful supervised classification algorithm but also identified the most suitable set of parameters between sagittal plane kinematics and sEMG. Our data also make it possible to identify the best set of variables extracted from the sEMG signal. For discriminating between FXS and CS individuals, all the proposed vectors allowed a high prediction accuracy (over 98.3%), but S2, using kinematic parameters, outperformed the other variables. For the classification of FXS-M within FXS-FM individuals the best prediction accuracy was achieved for S1 using all the sEMG parameters and the co-contraction as input features (97.0%).

The overall results of the present study suggest that by using sEMG standard parameters coupled with co-contraction indexes (S1), it is possible to classify subjects: FXS and CS as well as FXS-M and FXS-FM.

There is a paucity of information regarding specific physical therapy interventions that could impact the characteristic musculoskeletal alterations affecting FXS children. However, rehabilitation therapies for these children are mainly focused on the neurodevelopmental aspects and motor delay, and less attention is paid to actual musculoskeletal alterations [33]. Studies aiming at classifying the different gait motor control impairments associated with FXS could therefore help in overcoming this limitation. The results of our study showed the key role of joint kinematics and muscle activity in characterization and discrimination of the gait pattern of FXS children with respect to controls with a typical neurodevelopment. Furthermore, characteristic gait features were highlighted within FXS children carrying a full mutation versus those with a somatic mosaicism. This information could be used to stratify children according to the gait motor control impairments and associate this classification to specific physical therapy interventions.

The study has some limitations that should be acknowledged:

- The accuracy of several classifiers was higher for the test dataset than for the training one. There are many reasons for this result: the use of the regularization methods such as dropout, the fact that the validation set may include data that is easier to classify than the training one, the fact that the splitting between training and test sets may not be optimal [34];
- The low differences in the accuracy that might be caused by the small number of subjects forming the testing set (even though the ratio of correctly classified subjects might be the same, the percentage may differ due to the difference in the sets' sizes). Future works should consider increasing the studied population to strengthen the present results;
- The sample size within the study design, as previously mentioned;
- The lack of application of pre-processing procedures for dimensional data reduction (i.e., feature selection methods, PCA, methods for cross-validation and feature normalization) [9]. Future work should explore the possibility of reducing the set of parameters extracted from both kinematics and sEMG with the aim of enabling real-time application.

5. Conclusions

In summary, we can conclude that:

- An automatic recognition pipeline for gait disorders in FXS was proposed based on machine learning algorithms;
- Even though experimented in a limited cohort, this approach could offer a prompt assessment of gait motor control impairments in FXS subjects that could assist both in the diagnosis and treatment of this condition;
- The high accuracy revealed by kinematics, sEMG and co-contraction parameters support exploring their use for the development of assistive decision-making systems for the prescription of specific physical therapy programs for FXS individuals;
- This work has a translational impact since it provides objective and measurable parameters to evaluate important clinical aspects of the FXS phenotype that can lead to better outcomes for future clinical trials.

References

1. Coffee, B.; Keith, K.; Albizua, I.; Malone, T.; Mowrey, J.; Sherman, S.L.; Warren, S.T. Incidence of fragile X syndrome by newborn screening for methylated FMR1 DNA. *Am. J. Hum. Genet.* **2009**, *85*, 503–514. [CrossRef] [PubMed]
2. Hagerman, R.J.; Protic, D.; Rajaratnam, A.; Salcedo, M.; Aydin, E.Y.; Schneider, A. Fragile X-associated neuropsychiatric disorders (FXAND). *Front. Psychiatry* **2018**, *9*, 564. [CrossRef]
3. Johnson, K.; Herring, J.; Richstein, J. Fragile X premutation associated conditions (FXPAC). *Front. Pediatr.* **2020**, *8*, 266. [CrossRef]

4.　Hagerman, R.J.; Berry-Kravis, E.; Hazlett, H.C.; Bailey, D.B., Jr.; Moine, H.; Kooy, R.F.; Tassone, F.; Gantois, I.; Sonenberg, N.; Mandel, J.L.; et al. Fragile X syndrome. *Nat. Rev. Dis. Primers* **2017**, *3*, 17065. [CrossRef] [PubMed]

5.　Jiraanont, P.; Kumar, M.; Tang, H.-T.; Espinal, G.; Hagerman, P.J.; Hagerman, R.J.; Chutabhakdikul, N.; Tassone, F. Size and methylation mosaicism in males with Fragile X syndrome. *Expert Rev. Mol. Diagn.* **2017**, *17*, 1023–1032. [CrossRef]

6.　Jacquemont, S.; Hagerman, R.J.; Hagerman, P.J.; Leehey, M.A. Fragile-X syndrome and fragile X-associated tremor/ataxia syndrome: Two faces of FMR1. *Lancet Neurol.* **2007**, *6*, 45–55. [CrossRef]

7.　Sawacha, Z.; Spolaor, F.; Piątkowska, W.; Cibin, F.; Ciniglio, A.; Guiotto, A.; Ricca, M.; Polli, R.; Murgia, A. Feasibility and reliability assessment of video-based motion analysis and surface electromyography in children with fragile X during gait. *Sensors* **2021**, *21*, 4746. [CrossRef] [PubMed]

8.　Hallemans, A.; Van de Walle, P.; Wyers, L.; Verheyen, K.; Schoonjans, A.-S.; Desloovere, K.; Ceulemans, B. Clinical usefulness and challenges of instrumented motion analysis in patients with intellectual disabilities. *Gait Posture* **2019**, *71*, 105–115. [CrossRef] [PubMed]

9.　Figueiredo, J.; Santos, C.P.; Moreno, J.C. Automatic recognition of gait patterns in human motor disorders using machine learning: A review. *Med. Eng. Phys.* **2018**, *53*, 1–12. [CrossRef] [PubMed]

10.　Dobson, F.; Morris, M.; Baker, R.; Graham, H.K. Gait classification in children with cerebral palsy: A systematic review. *Gait Posture* **2007**, *25*, 140–152. [CrossRef]

11.　Kamruzzaman, J.; Begg, R.K. Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait. *IEEE Trans. Biomed. Eng.* **2007**, *53*, 2479–2490. [CrossRef]

12.　Zhang, Y.; Ma, Y. Application of supervised machine learning algorithms in the classification of sagittal gait patterns of cerebral palsy children with spastic diplegia. *Comput. Biol. Med.* **2019**, *106*, 33–39. [CrossRef]

13.　Mannini, A.; Trojaniello, D.; Cereatti, A.; Sabatini, A.M. A machine learning framework for gait classification using inertial sensors: Application to elderly, post-stroke and Huntington's disease patients. *Sensors* **2016**, *16*, 134. [CrossRef]

14.　Whitley, E.; Ball, J. Statistics review 4: Sample size calculations. *Crit. Care* **2002**, *6*, 335–341. [CrossRef]

15.　Nahhas, F.A.; Monroe, T.J.; Prior, T.W.; Botma, P.I.; Fang, J.; Snyder, P.J.; Talbott, S.L.; Feldman, G.L. Evaluation of the human fragile X mental retardation 1 polymerase chain reaction reagents to amplify the FMR1 gene: Testing in a clinical diagnostic laboratory. *Genet. Test. Mol. Biomark.* **2011**, *16*, 187–192. [CrossRef] [PubMed]

16.　Chen, L.; Hadd, A.G.; Sah, S.; Houghton, J.F.; Filipovic-Sadic, S.; Zhang, W.; Hagerman, P.J.; Tassone, F.; Latham, G.J. High-resolution methylation polymerase chain reaction for fragile X analysis: Evidence for novel FMR1 methylation patterns undetected in Southern blot analyses. *Genet. Med.* **2011**, *13*, 528–538. [CrossRef]

17.　Blumenstein, R. *Electrode Placement in EMG Biofeedback*; Williams & Wilkins: Philadelphia, PA, USA, 1980.

18.　Blanc, Y.; Dimanico, U. Electrode placement in surface electromyography (sEMG) "Minimal Crosstalk Area" (MCA). *Open Rehabil. J.* **2010**, *3*, 110–126. [CrossRef]

19.　Tomasi, C.; Kanade, T. Shape and motion from image streams: A factorization method. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 9795–9802. [CrossRef] [PubMed]

20.　Sawacha, Z.; Minelle, F.; Scarton, A.; Cobelli, C. Development of an automatic tracking software for out of water and underwater motion analysis. In Proceedings of the World Congress of Biomechanics (WCB), Boston, MA, USA, 6–11 July 2014.

21.　Sawacha, Z.; Spolaor, F.; Guarneri, G.; Contessa, P.; Carraro, E.; Venturin, A.; Avogaro, A.; Cobelli, C. Abnormal muscle activation during gait in diabetes patients with and without neuropathy. *Gait Posture* **2012**, *35*, 101–105. [CrossRef] [PubMed]

22.　Bonato, P.; D'Alessio, T.; Knaflitz, M. A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait. *IEEE Trans. Biomed. Eng.* **1998**, *45*, 287–299. [CrossRef]

23.　Spolaor, F.; Sawacha, Z.; Guarneri, G.; Del Din, S.; Avogaro, A.; Cobelli, C. Altered EMG patterns in diabetic neuropathic and not neuropathic patients during step ascending and descending. *J. Electromyogr. Kinesiol.* **2016**, *31*, 32–39. [CrossRef] [PubMed]

24.　McFadyen, B.J.; Winter, D.A. An integrated biomechanical analysis of normal stair ascent and descent. *J. Biomech.* **1988**, *21*, 733–744. [CrossRef]

25.　Sartori, M.; Maculan, M.; Pizzolato, C.; Reggiani, M.; Farina, D. Modeling and simulating the neuromuscular mechanisms regulating ankle and knee joint stiffness during human locomotion. *J. Neurophysiol.* **2015**, *114*, 2509–2527. [CrossRef]

26.　Wu, X.; Kumar, V.; Quinlan, J.R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Yu, P.S.; et al. Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **2008**, *14*, 1–37. [CrossRef]

27.　Horst, F.; Lapuschkin, S.; Samek, W.; Müller, K.-R.; Schöllhorn, W.I. Explaining the unique nature of individual gait patterns with deep learning. *Sci. Rep.* **2019**, *9*, 2391. [CrossRef] [PubMed]

28.　Jiménez-Grande, D.; Atashzar, S.F.; Martinez-Valdes, E.; De Nunzio, A.M.; Falla, D. Kinematic biomarkers of chronic neck pain measured during gait: A data-driven classification approach. *J. Biomech.* **2021**, *118*, 110190. [CrossRef]

29.　Balaji, E.; Brindha, D.; Balakrishnan, R. Supervised machine learning based gait classification system for early detection and stage classification of Parkinson's disease. *Appl. Soft Comput.* **2020**, *94*, 106494. [CrossRef]

30.　Darbandi, H.; Baniasad, M.; Baghdadi, S.; Khandan, A.; Vafaee, A.; Farahmand, F. Automatic classification of gait patterns in children with cerebral palsy using fuzzy clustering method. *Clin. Biomech.* **2020**, *73*, 189–194. [CrossRef]

31.　Armand, S.; Watelain, E.; Roux, E.; Mercier, M.; Lepoutre, F.-X. Linking clinical measurements and kinematic gait patterns of toe-walking using fuzzy decision trees. *Gait Posture* **2007**, *25*, 475–484. [CrossRef]

32.　Schöllhorn, W. Applications of artificial neural nets in clinical biomechanics. *Clin. Biomech.* **2004**, *19*, 876–898. [CrossRef]

33. Lieb-Lundell, C.C. Three faces of fragile, X. *Phys. Ther.* **2016**, *96*, 1782–1790. [CrossRef] [PubMed]
34. Gron, A. *Hands-On Machine Learning with Scikit-Learn and Tensor Flow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 1st ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2017.