Classification and clustering of low back pain patients by analyzing surface myoelectric signals

Srhoj-Egekher, Vedran

Doctoral thesis / Disertacija

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:168:794897

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-11-23**

Repository / Repozitorij:

FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repozitory

University of Zagreb

FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Vedran Srhoj-Egekher

CLASSIFICATION AND CLUSTERING OF LOW BACK PAIN PATIENTS BY ANALYZING SURFACE MYOELECTRIC SIGNALS

DOCTORAL THESIS

Zagreb, 2022

University of Zagreb

FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

Vedran Srhoj-Egekher

CLASSIFICATION AND CLUSTERING OF LOW BACK PAIN PATIENTS BY ANALYZING SURFACE MYOELECTRIC SIGNALS

DOCTORAL THESIS

Supervisor: Professor Mario Cifrek, PhD
Supervisor: Adjunct Assistant Professor Stanislav Peharec, PhD
Zagreb, 2022

Sveučilište u Zagrebu
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Vedran Srhoj-Egekher

KLASIFIKACIJA I GRUPIRANJE PACIJENATA S KRIŽOBOLJOM ANALIZOM POVRŠINSKIH MIOELEKTRIČNIH SIGNALA

DOKTORSKI RAD

Mentor: prof. dr. sc. Mario Cifrek
Mentor: nasl. doc. dr. sc. Stanislav Peharec
Zagreb, 2022.

Doctoral thesis is done at the University of Zagreb Faculty of Electrical Engineering and Computing, at the Department of Electronic Systems and Information Processing.

Supervisor: Professor Mario Cifrek, PhD
Supervisor: Adjunct Assistant Professor Stanislav Peharec, PhD

Doctoral thesis has: 135 pages

Doctoral thesis num.: _____

About the Supervisor (Professor Mario Cifrek, PhD)

Mario Cifrek was born in Varaždin in 1964. He received Dipl. Eng., M.Sc. and Ph.D. degrees in electrical engineering from the University of Zagreb, Faculty of Electrical Engineering and Computing (FER), Zagreb, Croatia, in 1987, 1992, and 1997, respectively. Since December 1987 he has worked in the Department of Electronic Systems and Information Processing, FER. In December 2012 he was promoted to Tenured Full Professor. During his work at FER, he participated in numerous Faculty working committees and served two terms as Vice Dean for Education from 2006 to 2010. From 2013 to 2017, he was a member of the Accreditation cCouncil of the Agency for Science and Higher Education, and since 2017 he has been a member of the Committee of the Accreditation Council for Follow-up within Reaccreditation Processes. Since 2013, he has been a member of the Scientific Field Committee for the Field of Electrical Engineering and Computing. He participated in seven scientific projects financed by the Ministry of Science, Education and Sports of the Republic of Croatia and in three projects of the Croatian Science Foundation. He led three bilateral Croatian-Chinese scientific and technological projects, two Proof-of-Concept (PoC) projects of the BICRO agency, one Horizon 2020 EIT Health RIS Innovation project, and one project under the call Strengthening capacities for research, development and innovation (STRIP). He is researcher on the project DATACROSS of the Centre of Research Excellence for Data Science and Advanced Cooperative Systems and has participated in the project “Application of Croatian Qualification Framework in the area of biomedical engineering”. He has led or participated in a number of collaborative projects with industry. He is a co-author of two books, 15 chapters in scientific books, more than 40 journal papers, 31 of them CC with more than 800 citations in WoS, and more than 180 conference papers in the field of biomedical engineering, sensors and electronic instrumentation. He participated in scientific or program committees at international conferences and also serves as a reviewer for various international journals. He received awards in national and international innovation exhibitions, and he is also active in science popularization. He is a member of professional associations IEEE (senior member), IFMBE, ESEM, Croatian Biomedical Engineering and Medical Physics Society and Croatian Society for Communications, Computing, Electronics, Measurement & Control. He is a member of the Croatian Academy of Engineering (HATZ), where he is a secretary of the Department of Systems and Cybernetics. He is member of Scientific Council for Technological Development of the Croatian Academy of Science and Arts (HAZU). He was awarded the bronze "Josip Lončar" medal (1988) for the graduate student score, silver “Josip Lončar” medals for outstanding master of science thesis (1992) and the doctoral dissertation (1997), gold “Josip Lončar” medal (2019) for a significant contribution to the Faculty by establishing procedures to assess, assure and improve the quality of education, and for raising the profile of the Faculty through their scholarly research and professional work,

Annual Award “Rikard Podhorsky” of the Croatian Academy of Engineering in 2014, the Annual award of the City of Zagreb in 2014, and a State award for science for the year 2020 in the field of technical sciences.

O mentoru (prof. dr. sc. Mario Cifrek)

Mario Cifrek je rođen 1964. godine u Varaždinu. Diplomirao je, magistrirao i doktorirao u polju elektrotehnike na Sveučilištu u Zagrebu Fakultetu elektrotehnike i računarstva (FER), 1987., 1992., odnosno 1997. godine. Od prosinca 1987. godine radi na Zavodu za elektroničke sustave i obradbu informacija FER a. U prosincu 2012. godine izabran je u zvanje redovitog profesora u trajnom zvanju. Tijekom rada na FER-u sudjelovao je u radu brojnih radnih tijela Fakulteta, a od 2006. do 2010. godine bio je prodekan za nastavu u dva mandata. Bio je član Akreditacijskog savjeta Agencije za znanost i visoko obrazovanje od 2013. do 2017. godine, a od 2017. godine je član Povjerenstva Akreditacijskog savjeta za naknadno praćenje u postupcima reakreditacije. Od 2013. godine član je Matičnog odbora za područje tehničkih znanosti, polja elektrotehnike i računarstva. Sudjelovao je na sedam znanstvenih projekata Ministarstva znanosti, obrazovanja i sporta Republike Hrvatske te na tri projekta Hrvatske zaklade za znanost, te tri međunarodna projekta. Vodio je tri bilateralna projekta znanstveno-tehnološke suradnje s NR Kinom, dva projekta iz Programa provjere inovativnog koncepta za znanstvenike i istraživače (PoC PUBLIC) agencije BICRO, projekt iz programa EIT Heath RIS Innovation te projekt iz programa Jačanje kapaciteta za istraživanje, razvoj i inovacije (IRI). Istraživač je na projektu DATACROSS Znanstvenog centra izvrsnosti za znanost o podatcima i kooperativne sustave te je sudjelovao u projektu Primjena hrvatskog kvalifikacijskog okvira u području biomedicinskog inženjerstva. Vodio je ili je bio suradnik na većem broju projekata suradnje s gospodarstvom. Suautor je dvije knjige, 15 poglavlja u znanstvenim knjigama, objavio je više od 40 radova u časopisima od čega 31 CC sa više od 800 citata u bazi WoS, te više od 180 radova na međunarodnim i domaćim znanstvenim skupovima u području biomedicinskog inženjerstva, senzora i elektroničke instrumentacije. Sudjelovao je u organizaciji međunarodnih znanstvenih skupova kao član znanstvenog ili programskog odbora. Recenzent je u većem broju međunarodnih znanstvenih časopisa. Dobitnik je nagrada na domaćim i međunarodnim izložbama inovacija i aktivan je u popularizaciji znanosti. Član je stručnih udruga IEEE (seniorski status), IFMBE, Hrvatskog društva za biomedicinsko inženjerstvo i medicinsku fiziku i Hrvatskog društva za komunikacije, računarstvo, elektroniku, mjerenja i automatiku. Redoviti je član Akademije tehničkih znanosti Hrvatske (HATZ) gdje je tajnik Odjela sustava i kibernetike. Član je Znanstvenog vijeća za tehnološki razvoj Hrvatske akademije znanost i umjetnosti (HAZU). Dodijeljena mu je brončana plaketa „Josip Lončar“ za uspjeh tijekom studija 1988. godine, srebrne plakete „Josip Lončar“ za značajan i uspješan magistarski rad, te za posebno uspješnu

doktorsku disertaciju 1992., odnosno 1997. godine, zlatna plaketa „Josip Lončar“ za značajan doprinos Fakultetu uvođenjem postupaka vrednovanja, osiguravanja i unapređenja kvalitete nastave te za podizanje prepoznatljivosti Fakulteta svojim znanstvenoistraživačkim i stručnim radom, godišnja nagrada „Rikard Podhorsky“ Akademije tehničkih znanosti Hrvatske za 2014. godinu, Nagrada grada Zagreba u 2014. godini te Državna nagrada za znanost za 2020. godinu za područje tehničkih znanosti.

About the Supervisor (Adj. Asst. Prof. Stanislav Peharec, PhD)

Stanislav Peharec was born in Pula in 1963. He received B.Sc. degree in somatopedy from the University in Belgrade, Faculty for Special Education and Rehabilitation, in 1985. He received M.Sc. degree from the Faculty of Kinesiology in Zagreb in 2000, and Ph.D. degree in the field of physical medicine and rehabilitation from the Faculty of Medicine of the University of Rijeka in 2007. After mandatory internship in the General Hospital in Pula, he started his independent business and founded the Polyclinic for Physical Medicine and Medical Rehabilitation in Pula. As of 2018 he has been part-time employed at University of Rijeka Faculty of Health Studies, filed of physiotherapy. As of 2018, he is also part-time employed at the Faculty of Medicine of the University of Rijeka. In January 2011 he was elected as assistant professor of the Faculty of Medicine of the University of Rijeka. From 2005 till 2013 he was external associate and lecturer of the Biomechanics course at Faculty of Medicine of the University of Rijeka. He participated in six scientific projects of Ministry of Science and Education of the Republic of Croatia. He published several papers in scientific journals and participated in several international and domestic scientific conferences in the field of physical medicine, rehabilitation, biomechanics and sport. Stanislav Peharec is a member of professional associations CROBE-MPS (Croatian Biomedical Engineering and Medical Physics Society) and Croatian Society for Vertebrology. He participated in scientific committees at international conferences. In his work, he contributed to the advancement of physical medicine and rehabilitation especially in sports as well as the development of biomechanics in sport and rehabilitation. He was awarded state medal "Franjo Bučar" as a member of national team that won gold Olympic medal in Athens. He is a collaborator of numerous top sports organizations for the needs of functional diagnostics, prevention and rehabilitation of injuries. He was awarded with several medals for innovation.

O mentoru (doc. dr. sc. Stanislav Peharec)

Stanislav Peharec je rođen 1963. godine u Puli. Diplomirao je 1985. godine na Defektološkom Fakultetu Sveučilišta u Beogradu, smjer Somatopedija. Magistrirao je 2000. godine na Kineziološkom Fakultetu u Zagrebu i doktorirao 2007. godine na Medicinskom Fakultetu Sveučilišta u Rijeci u polju Fizikalne medicine i rehabilitacije. Nakon obveznog staža u Općoj bolnici u Puli započinje samostalni rad te osniva Polikliniku za fizikalnu medicinu i medicinsku rehabilitaciju u Puli. Od 2018. godine u nepunom radnom odnosu zaposlen je na Fakultetu zdravstvenih studija Sveučilišta u Rijeci, smjer fizioterapija. Od 2018., u nepunom radnom vremenu zaposlen je na Medicinskom fakultetu Sveučilišta u Puli. Od 2015. godine je znanstveni

suradnik i predavač na doktorskom studiju Kineziološkog fakulteta Sveučilišta u Zagrebu. U siječnju 2011. godine izabran je u zvanje naslovnog docenta Medicinskog Fakulteta u Rijeci. Od 2005. do 2013. godine bio je vanjski suradnik i predavač kolegija Biomehanika na studiju fizioterapija Medicinskog fakulteta Sveučilišta u Rijeci. Sudjelovao je na šest znanstvenih projekata Ministarstva znanosti, obrazovanja i sporta Republike Hrvatske. Objavio je više radova u znanstvenim časopisima te sudjelovao na više međunarodnih i domaćih znanstvenih skupova u području fizikalne medicine, rehabilitacije, biomehanike i sporta. Stanislav Peharec član je stručnih udruga HDBIMF (Hrvatsko društvo za biomedicinsko inženjerstvo i medicinsku fiziku) i Hrvatskog vertebrološkog društva. Sudjelovao je u organizaciji znanstvenih skupova kao član znanstvenog odbora. Svojim radom pridonio je unaprjeđenju fizikalne medicine i rehabilitacije posebice u sportu kao i razvoju biomehanike u sportu i rehabilitaciji. Dobitnik je državne nagrade Franjo Bučar kao član reprezentacije koja je osvojila zlatnu medalju na Olimpijskim igrama u Ateni. Suradnik je brojnih vrhunskih sportskih kolektiva za potrebe funkcionalne dijagnostike, prevencije i rehabilitacije ozljeda. Dobitnik je više nagrada na području inovacija.

We are the extension of our parents. For most of the things we know and the person we became, we can thank our parents who put incredible love and hopes into us.

Thank you mama Slavica for your love and support all these years, in good and bad times. I hope your heart is now filled with pride.

Dear dad Neven, sorry for not giving you the joy of this moment while you were still among us. I hope you are watching us from above with peace and joy in your heart.

Dear grandpa Tomislav, this is for you also. You taught me the first Latin sayings and I remember the thrill when you allowed me to play on your old typing machine as a kid.

Abstract

Low back pain (LBP) is a global health problem phenomenon. Most patients are categorized as non-specific, thus requiring an individualized approach which still poses a major challenge. In this doctoral research, sEMG recordings from two pairs of lumbar muscle sites were collected during an isometric trunk extension exercise. Ninety-one subjects were included in the study; 29 patients with non-specific chronic LBP (CLBP), 25 patients with radiculopathy (RLBP), and 37 control healthy subjects (HS). Six best-performing time-domain raw features were employed to model contextual secondary feature groups. Neuromuscular LBP characteristics were described with coordination, co-activation, trends, and frequency-based fatigue measures. Altogether, a set of 327 secondary features was created where inputs into the classification models were further refined by employing neighborhood component analysis (NCA). NCA effectively reduced the number of features (<20 components), alongside preserving them in the original interpretable domain. A set of 23 different classifiers was employed and explored, resulting in classification accuracy of 0.94 for HS vs. LBP, 0.90 for HS vs. CLBP, 0.98 for HS vs. RLBP, and 0.89 for CLBP vs. RLBP differentiation. High median precision (0.97) and sensitivity (0.99) across all classifiers for HS vs. RLBP differentiation was obtained, with only three feature components utilized (out of 327). Support vector machines (SVM) and *k*-nearest neighbor (*k*NN) based classifiers consistently demonstrated best classification results. The obtained HS vs. CLBP differentiation model, based on thirteen NCA components, was further exploited to define smaller, more homogeneous CLBP subgroups. This was achieved by defining a set of characteristic CLBP patterns based on visual inspection of all CLBP subjects via the radial plot visual representation. Initially, five subgroups were isolated and validated against the unsupervised clustering methods, namely, self-organizing maps (SOM) and hierarchical clustering. Hierarchical clustering analysis (HCA) via the corresponding hierarchical tree representation (dendrogram) proved to well coincide with the initial visually isolated subgroups. In this way, overall seven CLBP groups were promoted, namely: *Regular CLBP pattern*, *Inhibited pattern*, *Type A coordination imbalance pattern*, *Type B coordination imbalance pattern*, *Exceeding coordination pattern*, *Near-regular pattern*, and *Co-activation imbalance pattern* with twenty-five out of twenty-nine subjects at hand being assigned into one of the seven groups. The remaining four subjects were labeled as non-specified due to not straightforwardly fitting into any of the elaborated subgroups. In order to provide a more formal procedure for assigning new CLBP patients into one of the seven subgroups, for each subgroup, an averaged template profile was constructed and a new distance metric *CLBP Pattern Distance* (CPD) was introduced, enabling the calculation for the closest CLBP pattern (subgroup) to be assigned to the respective patient. This template-matching procedure has been validated against the reference subgrouping result and a perfect match was confirmed. As a final step, for each CLBP subgroup, a corresponding

medical interpretation was proposed, thus enabling a direct connection between mathematical procedure and the deeper inference from the medical perspective to account for the complex multifactorial interaction among different neuromotor control, biomechanical, or even psychological aspects. The overall method, presented in this PhD thesis, has demonstrated the ability for successful patients subgrouping and subsequent more individualized rehabilitation treatments, backed with medical interpretations through contextual features modeling.

Keywords: classification, clustering, features modeling, interpretability, machine learning, non-specific chronic low back pain, patients subgrouping, radiculopathy, support vector machines, surface electromyography

Prošireni sažetak

Predstavlja se križobolja kao opći globalni zdravstveni problem sa svojim značajnim socio-ekonomskim utjecajem. Smatra se da više od 80% svjetske populacije iskusi neki oblik problema povezanog s križoboljom tijekom života. Pri tome, također, smatra se da je na godišnjoj razini križobolja prisutna kod 15-20% populacije [1, 2, 3]. Ovakve brojke, osim značajnog utjecaja na samo zdravlje pojedinaca te njihovu kvalitetu življenja i produktivnost, smještaju križobolju na treće mjesto po iznosu zdravstvenih troškova, odmah iza dijabetesa i koronarnih srčanih bolesti. Tako su primjerice, navedeni troškovi povezani s križoboljom, u SAD-u (period 1993.-2013. godina) bili procijenjeni na 90 milijardi USD godišnje [4, 5]. Jedan od ključnih izazova u medicinskoj dijagnostici i liječenju križobolje formuliran je činjenicom iz kliničke prakse, koja govori, da se oko 90% pacijenata s križoboljom ne može dijagnosticirati i kategorizirati na bolji način osim da se navede njihov problem s križoboljom navede kao nespecificiran [6]. To predstavlja značajna ograničenja u osiguravanju pravilnih rehabilitacijskih postupaka [7, 8]. Kao dio cjelokupnije slike višestrukih napora u proučavanju križobolje, daje se i pregled povijesti tehnika i korištenih pristupa u bavljenju križoboljom, kao i pregled najnovijih dostignuća i trenutnih dominantnih pravaca u istraživanju križobolje. Pri tome se uočava da je površinska elektromiografija (sEMG) i dalje dominantna tehnika kako zbog svoje neinvazivnosti, tako i zbog relativne pristupačnosti kod same primjene, a nadasve zbog korisnih uvida u funkcionalne aspekte motoričke kontrole i povezanu biomehaniku [9, 10, 11]. U tom kontekstu, uobičajeno se uz mioelektričnu aktivnost prate i ne-električne veličine vezane uz kinematiku [12]. Posebni naglasak u provedenim istraživanjima, bilo kao dio općenitog proučavanja značajki dobivenih sEMG-om, bilo u konkretnom proučavanju križobolje, pridaje se matematičkim postupcima i tehnikama obrade, analize i klasifikacije, pri čemu se pokazuje da su dvije dominantne klasifikacijske tehnike temeljene na linearnoj diskriminantnoj analizi (LDA) i metodi potpornih vektora (SVM). U novije vrijeme postoje također pokušaji primjene naprednijih neuronskih mreža i metoda dubokog učenja, ali takve metode zasad nisu poprimile širu upotrebu, vjerojatno i zbog određenih ograničenja u samog mogućnosti interpretacije i tumačenja tako dobivenih rezultata [13, 14]. Navedena istraživanja, koja su i većinom temeljena na dinamičkom izvođenju pokreta (za razliku od statičkih vježbi, tj. kontrakcija), prijavljuju klasifikacijske rezultate diferencijacije zdravih ispitanika od ispitanika s križoboljom u rasponu točnosti od 85%-95% [15, 16, 17, 18, 19, 20, 21, 22, 23].

Nadalje, predstavljaju se glavni izazovi u ovom području potaknuti nemogućnošću jasnog povezivanja trenutnih teorija, koje objašnjavaju nastanak i manifestaciju križobolje, s rehabilitacijskim postupcima. Sugerira se da se nemogućnost pronalaženja predvidljivih ili uspješnih rehabilitacijskih postupaka najčešće objašnjava neuspjehom u razvrstavanju pacijenata u podskupine [7, 8, 24, 25], pri čemu nije moguće jasno otkriti koji čimbenici su izravno povezani i

sa samom pojavom križobolje.

Analizirajući radove i rezultate istraživanja iz navedenog područja, može se uočiti neujednačenost prikazanih rezultata istraživanja od strane različitih autora, s često izravno proturječnim zaključcima. Opravdanje za takve nedosljednosti se djelomično može pronaći u složenosti problema koji uključuje različite anatomske, fiziološke, motoričke, neurološke, biomehaničke, pa čak i psihološke čimbenike, koji onda međusobnom interakcijom stvaraju složeni sustav koji se ne može opisati jednostavnim procesima ili analizirati pomoću jednostavnih postupaka [9, 26, 27].

U prošlosti se dovodila u pitanje općenita ispravnost kliničkih postupaka u dijagnosticiranju i klasifikaciji križobolje. Tako se u radu Bogduka i McGuirka [28] tvrdi da nema dovoljno dokaza koji bi potvrdili da bilo koji određeni klinički znak, ili kombinacija znakova, pronađena tijekom procesa pregleda, omogućava ispravnu ili pouzdanu dijagnozu. Tradicionalne teorije i modeli kao što su *prilagodba na bol* i *bol-spazam-bol* (tzv. začarani krug), uz konstataciju nemogućnosti jasnog povezivanja teorija s rehabilitacijskim postupcima, dovele su do nužnosti definiranje novih teorija. Takve nove teorije uzimaju u obzir različite promjene koje se javljaju kao dio motoričke kontrole i kontrole boli, a s namjerom da daju konzistentniji smjer za razvoj i poboljšanje kliničkih postupaka. Aktualno dobro prihvaćena teorija i pripadni model je tzv. *motorička prilagodba na bol*, predstavljen od strane Paula W. Hodgesa [7].

U svjetlu navedenog, kao prvi korak, za dublje razumijevanje ove problematike, ovaj doktorski rad ima za cilj istražiti općenite mogućnosti i dosadašnja postignuća razdvajanja ispitanika s patologijom povezanom s križoboljom od zdravih ispitanika. Drugo, cilj je identificirati i izdvojiti one parametre (čimbenike, značajke, obilježja) koji najviše pridonose ovom razlikovanju, s tendencijom da se takvi čimbenici prevedu u razumljivije (opisne) značajke koje bi poželjno mogle imati primjenu i kao dio stvarne kliničke prakse.

Obzirom na tendenciju da se prvenstveno istraže mogućnosti i predlože modeli značajki, za opis križobolje pomoću sEMG-a, predložen je takav protokol prikupljanja podataka (snimanja mioelektrične aktivnosti) te analize, klasifikacije i grupiranja, koji uklanja sve one postavke koje bi mogle pridonijeti pozitivnoj pristranosti dobivenih rezultata. Stoga, predloženi postav je sljedeći: klasifikacije i grupiranje se temelje samo na značajkama sEMG-a, za vježbu je odabrana jednostavna izometrička (statička) vježba izdržaja trupa, nadalje, koriste se samo četiri sEMG kanala te zaključno, za tehnike klasifikacije i grupiranja se koriste predefinirane metode bez dodatnog podešavanja parametara. Shodno tome, predloženi postav ima tendenciju da rezultate tretira s većim pouzdanjem prilikom predlaganja rješenja za problem podgrupiranja ispitanika s križoboljom.

Prvi konkretni korak, u omogućavanju ovog istraživanja, jest prikupljanje sEMG snimaka, kako zdravih (kontrolna skupina), tako i ispitanika s križoboljom (pacijenti). Uz pomoć medicinskog stručnjaka, definirane su tri različite kohorte muških ispitanika, što je rezultiralo s tri

skupine: zdravi ispitanici (HS), ispitanici s nespecificiranom (kroničnom) križoboljom (CLPB) i ispitanici s radikulopatijom (RLBP), pri tome navedene skupine se sastoje od trideset sedam, dvadeset i devet, te dvadeset i pet ispitanika, redom.

Kao dio protokola mjerenja, križobolja je inducirana izvođenjem izometričke vježbe ekstenzije trupa u konfiguraciji rimske klupe, slijedeći jasno definirane kriterije mjernog protokola kojeg su se ispitanici trebali pridržavati, uz podršku medicinskog osoblja. Odgovarajuća vježba i sEMG mjerenje zaustavljeni su kada ispitanik više nije mogao održavati položaj ekstenzije trupa, ili je dosegnuto vremensko ograničenje trajanja vježbe od maksimalnih 180 s. Snimke sEMG-a su dobivene pomoću elektromiografskog sustava s bežičnim površinskim EMG elektrodama, i to snimanjem lumbalne regije na dvije pozicije obostrano (ukupno četiri kanala). Navedeno je rezultiralo mioelektričnom aktivnošću zabilježenom u gornjem dijelu lumbalnog područja (ULES) i donjem dijelu lumbalnog područja (LLES), konkretno, uz položaje kralježaka L1-L2 i L4-L5, slijedno.

Nakon toga, dobivene sEMG snimke, uzorkovane na 1000 Hz, su dalje matematički obrađene koristeći tehnike obradbe signala, pri čemu je inicijalno izračunato i analizirano sljedećih četrnaest jednostavnih sEMG značajki (tj. *primarne značajke*): broj prolazaka kroz nulu (ZC), promjena nagiba signala (SSC), Willisonova amplituda (WAMP), srednja apsolutna vrijednost (MAV), integrirani EMG (IEMG), varijanca (VAR), korijen srednjih kvadratnih vrijednosti (RMS), duljina valnog oblika (WL), log detektor (LD), kurtozis ili mjera izbočenosti/spljoštenosti (KURT), koeficijent asimetrije (SKEW), permutacijska entropija (PE), frekvencija medijana spektra snage (MDF) i relativna promjena varijance (RVD).

U sljedećem koraku pristupilo se kreiranju novih složenijih značajki na temelju prethodno izračunatih primarnih značajki. Obrazloženje za uvođenje takve, dodatne nove, konstrukcije leži u namjeri da se osiguraju funkcionalne i opisne značajke koje odražavaju konkretne izazove povezane s problemom križobolje, a sve to uz mogućnost interpretacije rezultata. Jedan od glavnih zadataka bilo je omogućiti detektiranje pacijenata s križoboljom i posljedično osigurati točnu i konzistentnu diferencijaciju pacijenata unutar same grupe ispitanika s križoboljom. Naime, navedeno i dan danas predstavlja jedan od glavnih izazova u radu s pacijentima i provođenju rehabilitacijskih tretmana [8, 9, 12, 25]. Slijedom navedenog, primarne značajke su iskorištene za kreiranje navedenih složenijih značajki (tj. *sekundarnih značajki*), koje u kombinaciji s razumijevanjem problematike i domenskim znanjem iz područja križobolje rezultiraju novim kontekstualnim značajkama. Navedene značajke su organizirane u sljedeće grupe: Koordinacijske mjere, Koaktivacijsko poklapanje, Trendovi, Frekvencijske mjere zamora. Pri tome, za koordinaciju nam je bitna usklađenost između lijeve i desne strane ULES i LLES područja, zasebno. Kod koaktivacijskog poklapanja se promatraju sve četiri mišićne pozicije te se prati njihova zajednička usklađenost (sinkronizacija). Trendovi prate vremenske promjene kako unutar mišića (tj. mišićne pozicije) s početka i kraja kontrakcije, tako i relativne promjene

između mišićnih parova bilateralno i ipsilateralno, i to promatranjem kroz sve odabrane primarne značajke. Završno, kao zasebna skupina, prate se frekvencijske mjere umora izračunati na temelju frekvencijske značajke MDF te pripadnih regresijskih pravaca koji prate promjenu MDF u vremenu.

Navedeni postupak rezultirao je konstruiranjem ukupno 727 značajki, pri čemu, jedna od značajki je i računalno određeno vrijeme korisnog dijela sEMG snimke definirano početkom i završetkom isometričke kontrakcije. Navedena značajka je implicitno prikupljeni parametar inherentan samoj sEMG snimki te, iako po prirodi ne-električna veličina, prikljuičen je navedenom ukupnom skupu značajki.

Dosad prikazani postupak uključivao je svih 14 primarnih značajki koje su rezultirale navedenim velikim brojem sekundarnih značajki. Kao sljedeći korak, zadatak je bio odabrati podskup primarnih značajki za koje smatramo da pojedinačno najviše pridonose konačnoj točnosti klasifikacijskih modela. U tu svrhu provedena je statistička i korelacijska analiza pri čemu su detektirane pojedine podskupine značajki koje su pokazale viši stupanj međusobne koreliranosti, i to MAV, IEMG, LD, RMS, VAR kao jedna podskupina, te SSC i PE kao druga.

U narednom koraku dobivena saznanja su iskorištena na način da je kreirano nekoliko usporednih (pod)skupova sekundarnih značajki s ciljem usporedbe utjecaja prethodnog odabira različitih podskupova primarnih značajki na krajnje rezultate klasifikacije. Pri tome, kao dio postupka izdvajanja najboljih sekundarnih značajki (*engl. feature selection*), iz većeg skupa sekundarnih značajki, uveden je postupak *analize komponenti metodom susjedstva* (*engl. neighborhood component analysis*, NCA) kao glavne korištene tehnike u tu svrhu u ovom doktorskom radu. Ovaj postupak omogućuje značajno smanjenje broja značajki kojima se modelira problem, zadržavajući značajke, odnosno kontekstualne informacije, u izvornoj domeni.

Kao dio općenitog pristupa provjere odabira značajki, i pripadnih modela, korišten je klasifikacijski postupak s klasifikacijskom točnošću kao kriterijem validacije. Pri čemu je namjera bila detektirati, odnosno uspostaviti smisleni skup sekundarnih značajki koje se mogu povezati sa stvarnim zdravstvenim stanjem ispitanika (tj. kontekstualizirati), izraženim kroz mioelektričnu aktivnost snimljenu pomoću sEMG-a.

U sklopu spomenutih klasifikacijskih postupaka i modela korišteni su sljedeći klasifikatori: stabla odlučivanja (DT) s finim, srednjim i grubim razdiobama; diskriminantna analiza: LDA i kvadratna diskriminantna analiza (QDA); logistička regresija (LR); SVM s različitim tipovima jezgri (linearne, kvadratne, kubične, Gaussove) i s različitim skalama za Gaussove jezgre (fina, srednja i gruba); metoda najbližih susjeda (kNN) s finom, srednjom i grubom razlučivošću; zbirni klasifikatori temeljeni na *bagging* i *boosting* algoritmima te tehnikama slučajnih podprostora *engl. random subspace*. Sveukupno dvadeset i tri klasifikatora i pripadnih varijanti s predefiniranim postavkama u programskom paketu Matlab (R2020b), The Mathworks, Inc. Uz navedene klasifikatore, dostupni skup podataka (ispitanika) se sastojao od devedeset i jednog

ispitanika, podijeljenih u tri skupine (HS, CLBP, RLBP), s udjelom od 40,7%, 31,9%, odnosno 27,4% među skupinama. Demografska i biometrijska statistika nije ukazala na ikakva značajna odstupanja ili razlike među skupinama, obzirom na dob, visinu, težinu i BMI (Tablica 3.1).

Prvi korak validacije odabira (primarnih) značajki je proveden na način da je izračunat skup sekundarnih značajki temeljem samo na jednoj primarnoj značajki. Na taj način je izračunato 14 zasebnih usporednih skupova sekundarnih značajki. Zatim je nad svakim od navedenih zasebnih skupova provedena NCA procedura odabira podskupa sekundarnih značajki (uz inherentno reduciranje broja značajki). Za svaki od 14 rezultirajućih skupova provedena je klasifikacija s 23 različita klasifikatora uz unakrsnu provjeru (*engl. cross-validation*) s deset preklapanja. Pri tome, radilo se o klasifikaciji s dvije klase, HS (kao cjelovita skupina zdravih ispitanika), te LBP skupina koje se sastojala od CLBP i RLBP skupine pacijenata zajedno. Za dobiveni skup rezultata je napravljena rang ljestvica predstavljajući prosječni poredak uspješnosti klasifikacije za svaku od 14 primarnih značajki u 23 različite klasifikacije. Rezultati su pokazali sljedeći redoslijed prema tako definiranom kriteriju: SSC, PE, WAMP, VAR, ZC, RVD, RMS, LD, MDF, WL, MAV, IEMG, KURT, SKEW, sa sljedećim prosječnim rangovima 2,38, 2,62, 4,10, 5,05, 6,17, 6,29, 7,40, 8,19, 9,45, 10,07, 10,45, 10,57, 10,71, 11,55, redom. Primjećuje se da su SSC i PE pokazali međusobno slične i u usporedbi prema ostalim značajkama, bitno bolje pojedinačne rezultate. Uzimajući u obzir rezultate pojedinačnih klasifikacija, kao i uvide iz prethodne statističke i korelacijske analize, kao optimalan podskup primarnih značajki odabrano je sljedećih (prvih) šest: SSC, WAMP, VAR, ZC, RVD + PE, gdje je PE uključen u skup unatoč detektiranoj korelaciji s SSC, ali njegova pojedinačni doprinos, kao druge najbolje rangirane značajke, nije mogao biti zanemaren.

Kao sljedeći korak validacije, napravljena je usporedba klasifikacijskih rezultata (HS vs. LBP) između nekoliko različitih skupova sekundarnih značajki na način da su odabrani početno različiti podskupovi primarnih značajki, i to primjerice: (a) SSC, WAMP, VAR, ZC, RVD, SSC; (b) WAMP, VAR, ZC, RVD + PE; (c) Hudginov vektor (MAV, WL, ZC, SSC); (d) Du-ov vektor (IEMG, VAR, WL, ZC, SSC, WAMP), (e) vektor svih nekoreliranih primarnih značajki ZC, WAMP, VAR, WL, KURT, SKEW, PE, MDF, RVD s PE umjesto SSC. Na temelju kojih su izračunati sekundarni skupovi značajki. Pokazuje se da nekorelirane značajke (e) i najboljih šest značajki (b) rezultiraju s najvišim klasifikacijskim točnostima 0,95 i 0,94, redom, i to za nekoliko različitih klasifikatora. Dodatno, rezultirajući broj NCA komponenti za navedene skupove (e) i (b) je bio 29 i 20, redom. Skup (b) je u daljnjem radu odabran kao optimalan obzirom na klasifikacijsku točnost i broj rezultirajućih NCA komponenti.

Nadalje, kako bi se omogućio uvid u to koja kontekstualna skupina značajki (koordinacija, koaktivacija, trendovi i frekvencijske mjere zamora) pojedinačno najviše doprinosi rezultatima točnosti klasifikacije, i u kojoj mjeri, navedene skupine su dodatno i međusobno uspoređene. Pri tome se pokazalo da koordinacijske i koaktivacijske značajke pojedinačno najviše dopri-

nose točnosti klasifikacije, odnosno diferencijacije između HS i LBP grupe ispitanika. Za koordinaciju, koaktivaciju, trendove, frekvencijske mjere zamora, te pomoćnu grupu maksimalnih vrijednosti primarnih značajki, najbolji rezultati za točnost te pripadni klasifikator su redom sljedeći: 0,91 (kNN), 0,88 (SVM), 0,81 (SVM), 0,80 (QDA), 0,81 (kNN), pri čemu je za potpuni skup koji objedinjava sve navedene sekundarne grupe ta točnost (prethodno već spomenuta) bila 0,94 (SVM).

Konačno, kao dio postupka klasifikacijske verifikacije, izvršen je niz klasifikacija s dvije klase s ciljem diferencijacije između tri grupe ispitanika (HS, CLBP, RLBP), i to između sljedećih klasifikacijskih parova: (I) HS naspram LBP, (II) HS naspram RLBP, (III) HS naspram CLBP, te (IV) CLBP naspram RLBP, kako je prikazano i s rezultatima u Tablici 5.5. Navedeni klasifikacijski parovi su pokazali vrlo visoku klasifikacijsku točnost za parove (I) i (II) s točnostima do 0,94 i 0,98, redom, te vrlo dobru klasifikacijsku točnost za parove (III) i (IV) s vrijednostima do 0,90 i 0,89, redom. Pri tome kao najbolji klasifikatori su se nametnuli: kvadratični i kubični SVM, kNN) s finom rezolucijom ($k = 1$), te metoda slučajnih podprostora temeljena na kNN) metodi. Dodatno, uz mjere klasifikacijske točnosti, za sva četiri para izračunate su i mjere preciznosti (PPV) i osjetljivosti (RC), pri čemu za klasifikator kNN) s finom rezolucijom, koji pojedinačno pokazuje ujednačeno najbolje rezultate, vrijednosti PPV za parove (I) do (IV) su redom 0,85, 0,97, 0,91, 0,92, dok su za RC redom 0,92, 1,00, 0,92, 0,87.

Prethodno dobiveni klasifikacijski rezultati, za dane parove ispitanika, su pokazali konzistentnost kroz nekoliko različitih validacijskih postava te su kao takvi naznačili određenu vjerodostojnost i mogućnost za primjenu u postupcima podgrupiranja CLBP ispitanika u manje homogene skupine. Naime, u ovom doktorskom radu, fokus je stavljen na pružanje uvida u mogućnost podgrupiranja CLBP pacijenata, koristeći podskup značajki odabranih u postupcima predloženog modeliranja kontekstualnih značajki križobolje. Glavni preduvjet je bio zadržati značajke u izvornoj domeni, što je i postignuto NCA procedurom (odjeljak 5.1.2). Nakon analiziranih rezultata klasifikacije (Tablica 5.5), potvrđeno je da odabrane značajke imaju vrlo dobru diskriminatornu snagu u smislu odvajanja zdravih ispitanika od pacijenata s križoboljom. Štoviše, uočeni su dobri rezultati i u daljnjem razlikovanju zdravih ispitanika od CLBP ispitanika, kao i diferencijacija između CLBP i RLBP skupina.

Za razliku od diferencijacije između HS i RLBP, koja je modelirana sa samo tri NCA komponente, diferencijacija između HS i CLBP uvela je više komponenti (trinaest) u model, što je poduprlo pretpostavku o složenijoj, odnosno raznovrsnijoj dinamici neuromuskularne i motoričke kontrole u pozadini. Ova složenost i raznolikost može se i lijepo vizualno prikazati. Naime, za usrednjene CLBP i HS profile, uočava se da se ne mogu statistički jasno diferencirati. Međutim, pretpostavka je da se takvo razlikovanje, i to s primarnom intencijom diferencijacije unutar same CLBP grupe, može ostvariti za manje homogene CLBP podskupine. Navedeno razmatranje je opet u skladu s dosadašnjim raspravama o višestrukim varijacijama kod ispi-

tanika s križoboljom i nužnosti podgrupiranja za bolje razumijevanje nespecifičnog CLBP-a [25, 29].

Na temelju danih komponenti značajki, za svakog CLBP ispitanika (pacijenta), kreiran je radijalni prikaz navedenih NCA komponenti u prikazu razapetog poligona (konture) te je takav radijalni CLBP profil vizualno pregledan za svakog od 29 CLBP ispitanika pojedinačno. Na temelju napravljenog vizualnog uvida i kontekstualnog domenskog znanja, izolirano je pet početnih obrazaca podskupina kao predložaka za podgrupiranje. Ovaj preliminarni postupak verificiran je formalnijim postupcima temeljenim na nenadziranim tehnikama grupiranja. Upotrijebljeni su SOM i HCA gdje se pokazalo da HCA daje dosljedne i interpretabilne rezultate grupiranja koji se mogu prilično dobro povezati s inicijalno odabranim kandidatima za svaku predloženu podskupinu. Dendrogram hijerarhijskog stabla dodatno je sugerirao neke dodatne grupe (klastere) koje treba uključiti u naš skup CLBP podgrupa, tako da je na kraju ukupno sedam podskupina uzeto u obzir: (I) *Standardni CLBP obrazac*, (II) *Inhibirani obrazac*, (III) *Obrazac s tipom A koordinacijske neravnoteže*, (IV) *Obrazac s tipom B koordinacijske neravnoteže*, (V) *Obrazac s prekomjernom koordinacijom*, (VI) *Blizu-standardni CLBP obrazac* i (VII) *Obrazac s koaktivacijskom neravnotežom*. U svrhu omogućavanja jednostavne metrike za kategoriziranje bilo kojeg CLBP profila pacijenata prema najsličnijoj podskupini, uvedena je nova metrika nazvana *Udaljenost CLBP uzorka* (CPD), pri čemu je kao kriterij, za pridjeljivanje određenoj podskupini, odabrana najmanja udaljenost između tog pojedinačnog profila i uprosječenog profila pojedine podskupine (prethodno izračunati uprosječeni profil za svaku od sedam podskupina na raspolaganju).

Kao dodatan doprinos podgrupiranju i interpretaciji rezultata, u analizu su uključeni i metapodaci (trajanje izometričke kontrakcije, indeks tjelesne mase, *Oswestry Disability Index* upitnik (ODQ) i *Roland-Morris Disability* upitnik (RDQ), Vizualna analogna skala (VAS), Lasegue-ov test (FFK)) pri čemu se u nekim segmentima mogu napraviti poveznice između podgrupa dobivenih vizualnim i HCA grupiranjem, s jedne strane, te vrijednostima iz ovog skupa metapodataka, s druge.

Kao rezultat postupka podgrupiranja, ukupno 25 od 29 ispitanika kategorizirano je u odgovarajuću podskupinu s određenom razinom sigurnosti. Iako je nakon podgrupiranja obrađeni CLBP obrazac karakteriziran s malim veličinama pojedine CLBP podskupine, mogu se izvući neka opća opažanja o distribuciji. Sedam ispitanika se nalazilo u podskupini (I) i četiri ispitanika u podskupini (VI), tako da bi ukupno 11 od 29 ($\sim$35-40%) ispitanika moglo biti predstavljeno kao dio standardnog uzorka u širem smislu (obzirom i da pripadaju istom segmentu hijerarhijskog stabla grupiranja). Druga najveća konsolidirana podskupina je ona za (II) kojoj je dodijeljeno 5 ispitanika, što predstavlja $\sim$15-20% veličine CLBP uzorka. Sljedeća najkonsolidiranija podskupina bila je ona za (III) s četiri dodijeljena ispitanika ($\sim$15%). Preostale tri podskupine (IV), (V) i (VII) imale su samo 2 predstavnika po podskupini, tj. <10%. Preostala 4

uzorka nisu se adekvatno mogla svrstati u nijednu od navedenih sedam podskupina te su i dalje ostavljeni s kategorizacijom kao *nespecificirani*.

Na temelju rezultata podgrupiranja i kreiranih CLBP profila, analizirani su predloženi modeli značajki za križobolju, s ciljem pružanja adekvatne medicinske interpretacije.

Kao dio interpretacije rezultata, nekoliko bitnih aspekata je naglašeno. Jedan od njih je vezan za značenje i doprinos primarne PE značajke (odjeljak 3.2.2.12). Naime, ovu primarnu značajku karakterizira sposobnost kvantificiranja razine nelinearnosti u signalu i same složenosti procesa u pozadini. U kontekstu zabilježene i analizirane mioelektrične aktivnosti, u ovom doktorskom radu se pretpostavlja da takva značajka, i odgovarajuće mjere koje se mogu posredstvom nje kvantificirati, pruža uvide u naprednije strategije neuralne motorike koje rezultiraju pojačanom (složenijom) koordinacijom i koaktivacijom. Složenost se može izraziti kao superpozicija različitih komponenti ili strategija motoričke kontrole, odnosno aktivacije ili deaktivacije određenih motoričkih jedinica, angažiranja dodatnih motoričkih jedinica u istoj regiji ili preraspodjele opterećenja na sinergističke mišiće. Mogućnost pružanja složenijih strategija motoričke kontrole se, prema dobivenih uvidima iz ove analize, općenito više odnosi na HS skupinu nego na CLBP i RLBP, iako pojedini CLBP uzorci kao iz podskupine (III) i (V) pokazuju upravo suprotno, tj. pojačanu koordinacijsku aktivnost primjerice u LLES regiji. Također, Za pojedine uzorke, predložena interpretacija bi mogla upućivati na to da je moguće prisutna veća mioelektrična aktivnost u ULES području, što bi bilo u koliziji s nekim općim zapažanjima u radovima drugih autora, gdje je općeprihvaćeno da je za CLBP skupinu mišićna aktivnost u LLES-u izraženija nego u ULES području [9, 30, 31]. Međutim, prikazano stajalište ne navodi eksplicitno da je mišićna aktivnost u području LLES manje izražena nego u području ULES. Interpretacija, predložena u ovom doktorskom radu, ističe samo da kompenzacijski mehanizmi u područjima LLES-a očito nisu bili dovoljni, pa su se pretpostavljeno uključile dodatne strategije preraspodjele i kompenzacije koje su uključivale i područje ULES-a.

Zaključno, predstavljeni postupci za kreiranje značajki i pripadnih kontekstualnih modela, provedbu klasifikacije u svrhu validacije modela, te primjenu postupaka grupiranja, rezultirali su novim pristupom u modeliranju problema križobolje pomoću sEMG značajki, rezultirajući sa sedam distinktnih CLBP podgrupa s predloženim interpretacijama utemeljenima na složenom medicinskom pogledu uzimanjem u obzir složenih neuromotoričkih, biomehaničkih, fizioloških, kao i psiholoških čimbenika. Navedena metoda posljedično može omogućiti individualiziraniji pristup tijekom medicinskih tretmana pacijenata s križoboljom u kliničkoj praksi, pri čemu se povećavaju izgledi za bolje rezultate rehabilitacijskih postupaka.

Ključne riječi: grupiranje, interpretabilnost, klasifikacija, nespecifična kronična križobolja, metoda potpornih vektora, modeliranje značajki, podgrupiranje pacijenata, površinska elektromiografija, radikulopatija, strojno učenje

Contents

Chapter 1

Introduction

In this chapter, as an introduction, low back pain (LBP) as a general global health problem is inaugurated with its significant socio-economic impact. One of the key challenges in medical diagnostics and treatments is formulated with a real-life fact that about 90% of patients, with LBP, are being no better diagnosed and categorized, but with stating that their LBP issues are *non-specific*. This poses significant limitations in ensuring proper rehabilitation procedures. The history of techniques and approaches in tackling LBP is presented with the most recent state-of-the-art advancements and current dominant directions in LBP research via sEMG.

1.1 LBP impact on society

Low back pain (LBP) has been clearly identified and confirmed as one of the major global public health problems for decades now, ranked as the number one cause of years lived with disability [1]. It is estimated that more than 80% of the population will experience some form of LBP-related problems during its lifetime, whereas the yearly prevalence of this problem is present among 15-20% of the population as reported by studies by Anderson *et al.* [2] and Breivik *et al.* [3] or as more recently reported, with the global yearly prevalence of over 500 million people at any time as stated by James *et al.* [1]. This phenomenon, taking large scale, not only affects the health, abilities, and productivity levels of the individuals but has tremendous socio-economic implications, consequently incurring significant costs. Dieleman *et al.* reported that low back and neck pain are the third costliest health-care related direct expenditures (approx. US $90 billion a year) in the USA (for the period 1996-2013) [4], following diabetes and ischemic heart diseases, where the indirect costs for the individuals and households are raising up to eight times compared to the direct health-care costs [5]. Similar studies were also conducted in some of the western and developed European countries (UK, Germany, France, Italy) with significant private or public health-care services expenditures [32, 33]. It is evident that despite very clear awareness of the problem stated, costs incurred, strong efforts, and much research

conducted with the aim to either prevent LBP or secure better LBP treatments and rehabilitation outcomes, many of the problems and challenges related to LBP are still present. Moreover, LBP implications on the public health and public health-care systems, when comparing the 1970s, 1980s, and today, are not reducing, but the implications have a trend of further increasing, where the modern way of living is also playing a significant role [4, 34].

1.2 Non-specific LBP

Another revealing insight is that, for about 90% of patients with LBP disorders, it can not be clearly stated the cause of the pain, thus such patients are being classified as *non-specific* LBP with treatments directed primarily on reducing pain itself and its consequences. Among patients with LBP disorders, 5-10% of patients can usually be related to radicular syndrome [6]. Such an approach, with no specification, cannot tackle the real causes of LBP problems and more subtle musculoskeletal changes with neuromuscular adaptations in behind [7]. Equally, guarantees of long-term positive outcomes for medical treatments and rehabilitation are thus less likely to hold. This led to hypothesizing that (non-specific) LBP patients are not a homogeneous group [8], but more likely a more individualized approach in diagnostics and rehabilitation is needed to account for diversity and complexity in behind [7, 8]. An additional challenge is a high level of reported ambiguities in understanding muscle recruitment mechanisms in LBP patients (as thoroughly summarized by Dieën *et al.* [9]), thus becoming the obstacle in supporting better treatment outcomes overall. In that course, considering and introducing multiple factors in terms of characteristics describing LBP could enable identifying more specific motor control patterns, potentially leading to a more homogeneous subgrouping of LBP patients. This would be of high importance.

1.3 State-of-the-art

There has been a history of different methods proposed and techniques employed to assess, predict or discriminate LBP in patients, as summarized in several review studies [9, 10, 11]. Some of these methods are relying on different imaging techniques, like highly invasive X-ray or Computed Tomography (CT) [35], less invasive ultrasound imaging [36], or non-invasive, but expensive MRI [37]. Such imaging techniques can reveal locations of structural and morphological changes (or damage) in the musculoskeletal system, however not always being able to pinpoint the real causes of pain. Moreover, some reports state that in up to 85% of cases no radiologic abnormality is evident [12]. Aside from detecting deep structural changes, in many cases it is of key interest to assess and recognize the effects of LBP on a person's ability to perform movements, resulting from motor-control impairments or bio-mechanical changes due

to LBP presence or history. Such insights consequently lead to better rehabilitation treatments and follow-ups. Kinematic parameters tracking is another, non-invasive, approach that enables such insights into bio-mechanical aspects of a person's movements influenced by LBP, by capturing altered gait patterns, impacted range of motion, or spine displacements with respect to the reference points [38, 39]. Nonetheless, the most commonly used approach to assess LBP in patients is through surface electromyography (sEMG). sEMG is a well-established technique that enables the non-invasive recording of muscle activity. The myoelectric activity reflects the neuromuscular adaptations in LBP patients, thus enabling more insights into motor control coordination and co-activation strategies among muscles. [9, 15, 30, 40, 41, 42].

Available cross-sectional studies tackling LBP detection or classification by means of sEMG only, were mostly focusing on separation between non-specific chronic LBP (CLBP) and healthy subject (HS) groups [15, 16, 17, 18, 19, 20, 23]. Interestingly, a recent review study by Tagliaferri *et al.* [11], on artificial intelligence (AI) and machine learning (ML) based LBP detection, reported only few sEMG-only based research papers to be considered with a "fair quality", i.e. with no shortcomings [21, 23], whereas the remaining ones were stated to be with "poor quality". Nonetheless, for the presented papers at hand, the classification accuracy for separating CLBP from HS was ranging between 80% in [20] to 98.04% in [23] (for one of the feature subsets).

Often, sEMG is combined with kinematic tracking to provide more controlled insights into the related muscle activation and movement patterns [30]. This enables expansion of the set of characteristics (beyond sEMG features only) that are subsequently employed for detection or discrimination of LBP patients [12, 43, 44]. Generally, it is observed that such studies, combining sEMG and kinematic parameters, were resulting in better classification accuracy results compared to sEMG-only ones. These studies were also, among the rare ones, to tackle further discrimination within CLBP groups. Dankaerts *et al.* [12] reported overall classification success in discriminating healthy controls and two clinical subgroups (Flexion Pattern, FP, and Active Extension Pattern, AEP, patients) with 96.4% accuracy, however, the accuracy of 84.1%, if only sEMG features were utilized. Olugbade *et al.* employed classifiers in recognizing different LBP levels (low and high pain levels) in patients, with the accuracy of 94% for the best-performing support vector machines (SVM) classification model tested (random forests, RF, were also employed). Another study, moreover utilizing sEMG features only, employed decision trees in differentiating CLBP and patients with radiculopathy (RLBP) from the HS group, with at best accuracy of 86.8% [22].

Furthermore, some of the studies tackled the prediction of rehabilitation outcomes. Liew *et al.* [44] employed three different binary classification models for differentiating between patients with current LBP and with LBP in remission, alongside a third control group, where the best-reported area under the receiver operator curve (AUC) of 96.7% was obtained for differen-

tiation between two LBP groups (current and remission). A similar cohort study was presented by Jiang *et al.* for discriminating the "responding" against the "non-responding" LBP patients' treatment outcomes [21], by utilizing sEMG only features, with a best-reported accuracy of 96.67%.

Dominant classification models employed in presented studies were based on discriminant analysis [15, 16, 17, 18, 19], where in more recent works, SVM classifiers are gaining more spotlight [21, 23, 43]. There have been also some attempts to apply deep learning methods in sEMG-based LBP recognition with convolutional neural networks (CNN) [13], but still without gaining more traction, although reporting a high average classification accuracy of 92.9%. Though, some more papers were investigating deep learning methods in forecasting sEMG feature for trunk muscle fatigue [14]. Or, a study for recognizing LBP individuals from healthy ones by employing (non-SEMG) kinematic parameters as input into deep learning neural network with long short-term memory (LSTM), and reporting both classification precision and recall up to 97.2%.

It is also worth noticing that in most cases those studies relying on the dynamic functional type of exercise tasks, were the ones also demonstrating the best classification results [18, 21, 23, 43, 44]. Where also, in many cases, these studies were collecting inputs from multiple muscles or muscle sites [12, 18, 44]. Some of the examples, where good classification results were obtained with fewer sEMG channels, were presented in [23, 43], where furthermore, a simple isometric Biering-Sorensen exercise [45] was employed in [13].

1.4 Organization of the text

The remaining of the PhD thesis is organized in the chapters as follows:

- •Problem definition (Chapter 2): gives medical background, sEMG characteristics and points out to major challenges pertained to LBP with setting research hypotheses
- •SEMG recording and features extraction (Chapter 3): describes the measurement protocol setup employed for sEMG signal acquisition and first-step primary (raw) features extraction
- •Contextual features modeling (Chapter 4): provides the rationale behind introducing contextual features and formalizes mathematical and algorithmic procedures for features calculation
- •Classification modeling and features selection (Chapter 5): establishes approach and techniques for both primary and secondary features analysis and selection, presents classifiers employed, and with the final step verifies the proposed modeling approach validity through the classification outcomes
- •Clustering and CLBP patients subgrouping (Chapter 6): provides deeper insight into the

CLBP patients group by proposing a subgrouping approach and validating detected LBP subgroup patterns by means of different clustering methods

•Discussion and interpretation (Chapter 7): here the overall presented methodology is discussed reflecting both the obtained feature modeling, classification, and subgrouping (clustering) results. The corresponding medical interpretation for detected CLBP subgroups is presented as well

•Contributions and conclusion (Chapter 8): the final chapter recaps the main contributions as part of this PhD thesis, alongside other improvements or other non-standard methods employed in CLBP classification and clustering

Chapter 2

Problem definition

This chapter draws main challenges in this area driven by the inability to clearly link current theories, which explain the origin and manifestation of low back pain, with rehabilitation procedures. It is suggested that the inability to find predictable or successful rehabilitation procedures is most often explained by the lack of success in classifying patients into subgroups [7, 8, 24, 25], where it is not possible to clearly detect which factors are directly related to the incidence of low back pain.

Analyzing the papers and research results from the given field, the inconsistency of the presented research results among different authors can be observed, with often directly contradictory conclusions. The justification for such inconsistencies can be found, partially, in the complexity of the problem involving various anatomical, physiological, motor, neurological, biomechanical, and even psychological, factors that interact creating a complex system that cannot be described by simple processes or analyzed by simple procedures [9, 26, 27].

In the past, the correctness of clinical procedures in diagnosing and classifying LBP has been questioned. One example is work by Bogduk and McGuirk [28] where it is argued that there was no sufficient evidence to confirm that any particular clinical sign or combination of signs, found during the examination process allows for a correct or reliable diagnosis. The inability to clearly connect theories with rehabilitation procedures led to new theories. Such new theories take into account various changes that occur in pain and motor control with the intention of providing direction for the development and improvement of clinical procedures [7].

As the first step, toward a deeper understanding of this issue, this PhD thesis aims to explore the general possibilities and so far achievements of separating subjects with the pathology associated with low back pain from healthy subjects. Secondly, the aim is to identify and single out those parameters (features, predictors) that contribute most to this distinction, with the desired tendency that such factors can be translated into better understandable (descriptive) features or factors that favorably could have an application in the actual clinical practice as well.

2.1 Medical background

2.1.1 Anatomy and muscles

One of the important steps in the process of monitoring and analyzing LBP is the selection of muscles, or muscle groups, from which myoelectric activity would be tracked and correlated with the associated neuromotor and biomechanical manifestations of LBP. For this purpose, it is necessary to define the anatomical locations where it will be best possible to record such muscle activity as directly as possible, and with as little impact from EMG activity crosstalk of neighboring muscles as possible. The review of the research papers shows the following muscle groups as those of interest in tracking the manifestations of LBP and accompanying adaptive, co-activating, and synergistic mechanisms:

- upper lumbar erector spinae, ULES
- lower lumbar erector spinae, LLES
- hip extensors
- abdominal muscles

The main muscle groups of interest are those related to the upper and lower lumbar regions (ULES and LLES). Abdominal muscles and hip extensors have only recently begun to arouse more interest in the study of LBP, and efforts are being made to more clearly define their role and contribution to the various co-activation and synergistic mechanisms that accompany the emergence of LBP [12, 30, 40]. Muscles anatomy in more details is shown for abdominal muscles Fig. 2.1 and back muscles Fig. 2.2.

The position of the electrode is usually determined by the position of the spine, bilaterally and symmetrically in relation to the vertical axis of the spine. In the spinal area (Fig. 2.3), the electrodes are placed in the range of the T9 area (thoracic vertebrae), over the lumbar vertebrae (L1-L5), down to the first sacral vertebrae (S1).

Muscle activity in the ULES area is most often recorded at the positions L1-L2 [30, 49], while in some works myoelectric activity in this area is also recorded at the position T9 [17]. Similar differences can be found in defining the LLES area of interest, where LLES measurements at L3 [17] position, L4 [30] position, and L5-S1 [49] position are also performed. In many cases, the L5 position is related to the measurement of *multifidus* muscle activity [12, 17, 31, 50] which belongs to the group of deep back muscles and as such cannot be measured directly (i.e. without crosstalk from surrounding muscles) when exploiting the surface methods such as sEMG.

It is interesting to analyze the results of research works related to the contributions of individual muscles, or their activity, as an indicator of the presence of LBP or as a source of features in the classification of subjects into diagnostic groups. The vast majority of papers confirm the role of the muscle *multifidus* (L5) as a muscle that carries essential information to detect the

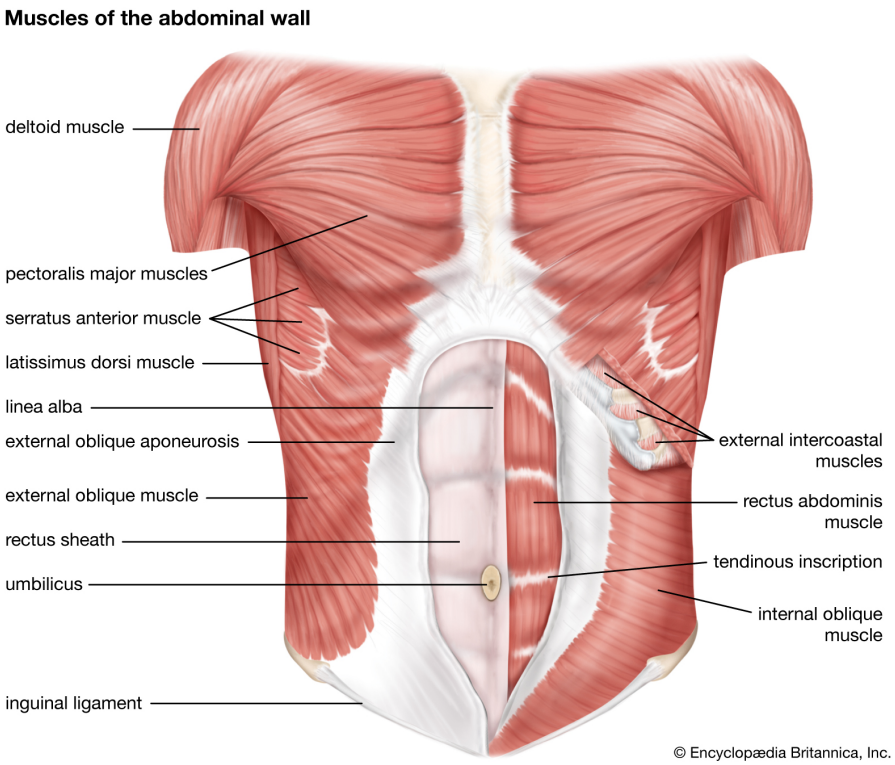


Figure 2.1: Muscles of the abdominal wall. Available online from Encyclopædia Britannica under citing terms [46].

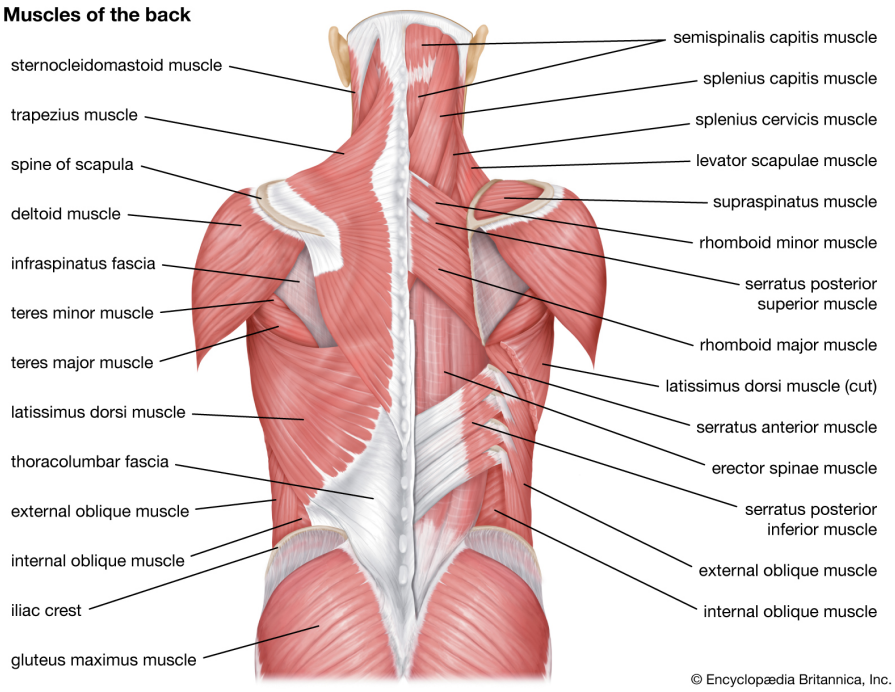


Figure 2.2: Muscles of the back. Available online from Encyclopædia Britannica under citing terms [47].

development of fatigue [31, 51]. It is also noted that it has an important discriminatory role in performing different exercises/movements among different groups of subjects with LBP [12].

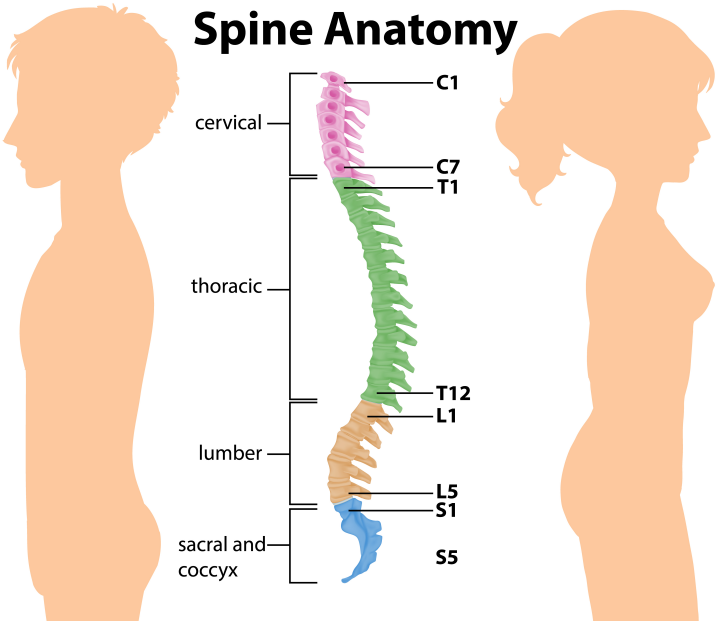


Figure 2.3: Visualization of the spine with vertebral distribution: cervical, thoracic, lumbar, sacral, and coccyx. Downloaded from Vecteezy and attributed to Spine Anatomy Vectors by Vecteezy under Free License terms [48].

Furthermore, it is stated that *multifidus* (L5) carries more significant information than L1 and L2 muscle sites [17, 31]. Although in some papers, somewhat contradictory, it is stated that in LBP subjects during flexion and endurance exercises, the highest muscle activity is recorded in the L2-L3 area [52]. The explanation of such results is possible only through a more comprehensive analysis of the respective results.

Muscles *gluteus maximus* and *m. biceps femoris* (one of the muscles of the posterior thigh of the upper leg (*hamstring*)), have certain correlations with the manifestations of LBP. For *m. gluteus maximus* LBP correlation with faster onset of fatigue [40], as well as shorter duration of muscle activation [53], is reported. For the *hamstring* muscles, certain correlations are observed between lumbar fatigue, inhibition of neural activation of quadriceps muscles (although without the appearance of fatigue), and, presumably, the consequent development of fatigue in the posterior thigh muscles. At the same time, in LBP subjects, these phenomena are somewhat even more pronounced [54]. Among other insights, for *hamstring* muscles, a significant correlation between the spectral indicators of fatigue and the time of fatigue onset is reported. Though, similar significance, but somewhat less convincing is also being stated for *m. biceps femoris*, but not for ULES and LLES muscle regions [49].

Abdominal muscles show a significant contribution to ensuring the stability of lumbar segments and contribute to the control of spinal movements, through co-activation protection mechanisms (antagonistic pair with muscle *m. multifidus*), which is especially important for CLBP patients. A special contribution to the stabilization of the lumbar region comes from the deep muscle *m. transversus abdominis* (TA) through its co-contraction and preparatory role in per-

forming the movement [55, 56]. The superficial abdominal muscle, *m. rectus abdominis* (RA), has a somewhat less effective role in stabilization which also participates as a co-activation pair with muscle *m. multifidus* (which is additionally pronounced in CLBP subjects), similar to the contribution of muscle *m. obliquus abdominis externus* (OE), where the total contribution of OE is further questioned [30, 56].

2.1.2 LBP patients differentiation

The main group of interest in most studies are the subjects with chronic LBP, as it can be confirmed from the review studies given by van Dieën *et al.* [9] and Geisser *et al.* [41]. There are several criteria for inclusion of subjects in this group, primarily related to the duration of LBP with localized pain in the lumbar area and with a clear mechanical basis of pain associated with posture and movement [12]. Although some papers include persistent or frequent pain of at least 12 months [31] or 6 months [17] as an inclusion condition for CLBP, most authors use persistent or frequent pain of at least 3 months duration as a criterion [12, 30, 50]. One of the main exclusion criteria is the subject's medical history, which includes spinal or pelvic surgeries [30, 31, 50], whereas in some papers such subjects with a medical history of spinal surgeries are included in the CLBP group [15]. Other exclusion criteria include: structural pathologies of the spine resulting in inflammation, tumors in the lumbar region, more severe spinal deformities, neurological or musculoskeletal disorders affecting gait, then diabetes, disc herniation, or subjects with radiculopathy [18, 30], where, for some authors, radiculopathy is not an exclusive criterion of [50].

The patients are considered with acute low back pain, if severe LBP problems were lasting no more than 4-8 weeks preceding the respective studies [18, 57, 58].

The control group, i.e. the group of subjects without LBP, includes those subjects without a history of back pain within the last 12 months [30, 50] or in some cases 24 months [12], without previous spine surgery [30, 50], without neurological or musculoskeletal disorders affecting gait [30]. In some papers, prior education on proper posture is also taken as a criterion [12].

Given the overview, it can be concluded that the criteria are not always clear or uniform, nonetheless, some basic conditions in terms of the presence of some form of LBP and its duration, as well as the history of surgery in the spine are common to all, or at least the vast majority of papers [9, 41].

By analyzing the sample sizes, i.e. the number of subjects (patients) in groups and subgroups (usually two: CLBP and control group), it can be seen that in relevant papers (by citation criterion) the total number of subjects is usually > 50, and usually closer to 100, where the number of subjects with LBP is, as a general rule, at least equal to the number of subjects in the control group, and often 2-3 times higher, taking into account the intention to further classify subjects from the LBP group into subgroups. Proper preparation and sizing of the sample are

deemed important for statistically relevant analysis [30, 49, 59] as well as the possibility of classifying subjects into subgroups.

2.1.3 Inducing LBP through exercise

The aim of the exercise is to bring the subject into a state in which the compensatory mechanisms will not be able to overcome the mechanisms of LBP emergence. This condition is generally achieved by some sort of activity or fatigue, that is suitable for the onset of LBP, which we then try to detect by employing muscle activity measurement procedures, such as sEMG. The main structures involved in compensatory mechanisms are: muscles (as the main stabilizers and load carriers), tendons, ligaments, joints, and fascia [26]. The onset of LBP, i.e. the inability of subjects to ensure stability and effective coordination between these structures and the neuromuscular system, will result in: changes in mobility, changes in the curvature of the spine, unadjusted fascia tension, improper joint function, decreased strength and endurance, muscle co-activation, muscle hypotonicity, tissue inflammation due to compression, etc. [7, 26, 54].

The types of exercises and activities, that we encounter in research on subjects with LBP, can be divided into several basic categories: (i) isometric endurance exercises (variants of Biering-Sørensen exercises or full trunk extension) [45, 54], ii) simple static positional exercises, with minimal or no load (sitting, standing, standing on one leg, an incline of $20^\circ - 30^\circ$, etc.) [9, 12, 26, 60] and (iii) dynamic exercises that include patterns of flexion and extension movements as well as walking [12, 30]. It can be observed that most EMG studies involve exercises in the state of rest, and only a few studies are based on isometric exercises and the maximum strength of performing an exercise or movement [9]. One of the reasons is that such simple exercises (at rest) are not subject to inaccuracies as exercises that require voluntary control and more significant use of force [12]. At the same time, such exercises enable the performance of movements that better imitate movements in normal daily activities [61]. An interesting result reported in the paper Luoto *et al.* [62] states that the exercise of standing on one leg shows the greatest sensitivity in distinguishing LBP subjects from the control group, where LBP subjects are characterized by poorer balance.

2.2 LBP mechanisms and models

One of the main challenges in analyzing LBP is understanding the mechanisms of pain and changes in the neuromuscular system that either supports the pain or find mechanisms in order to adapt so that an individual, with present tissue damage, can perform certain movements with minimal pain. Such an understanding would enable better detection of various LBP expressions and detection of LBP in the early stages, which would lead to better prevention, as well as help

in the rehabilitation processes. There are two classical models that have tried to explain these mechanisms, namely *pain-spasm-pain* (or the so-called "vicious cycle"), and *pain adaptation*, i.e. adjustment to pain [9]. As van Dieën *et al.* [9] presented, after analyzing thirty studies dealing with low back pain, neither of these two models could have fully explained the mechanisms of origin, manifestation, or adaptation to LBP, but moreover, often led to conflicting conclusions. Trying to better understand the mechanisms that accompany LBP, special attention was also paid to the study of *spinal stability*, by analyzing the role and interaction of passive and active stabilization structures [26]. Finally, the efforts to offer a comprehensive model, which would take into account the interdependence of a wider range of factors and try to explain the contradictions between individual models and research, resulted in the emergence of a new model called *motor adaptation to pain* [7, 27].

2.2.1 Pain-spasm-pain

The older *pain-spasm-pain* model assumes that the initial pain results in increased muscle activity in the painful area, i.e. in hyperactivity or spasm, which in turn leads to additional pain due to ischemia and the vascular system's inability to remove metabolites (e.g. lactate) which have been shown to have a strong irritant effect on pain receptors in muscle tissue [63, 64]. Although this approach, as Hodges *et al.* [7] stated, had initial support, two shortcomings undermine the validity of this approach. Firstly, although there is evidence of increased muscle activity, this phenomenon is not ubiquitous. Namely, many studies show a decrease in muscle activity [65] or report activity without change [66, 67]. Secondly, there have been clinical improvements without changes in muscle activity [68]. This suggests that clinical success is more related to psychological circumstances than to the rehabilitation of motor adaptation itself.

2.2.2 Pain adaptation

The second model, the so-called adaptation to pain, assumes that pain reduces muscle activation when muscles act as agonists or increases muscle activation when they are acting as antagonists. This adjustment results in a reduction in the speed and range of motion, which aims to prevent the mechanical infliction of pain in damaged tissues and their further damage [69]. Data supporting this theory can mostly be found in cases of acute pain caused by experiments. Thus, for example, agonist muscle activity is reduced during voluntary movements of the painful jaw [70], trunk [65], neck [71] and limbs [72], and a combination of decreased and increased muscle activity is recorded during dynamic leg movements [73]. However, although many studies support this model, they are not supported in all cases [7, 9].

2.2.3 Spinal stability

An additional circumstance that is taken into consideration, related to the model proposal and explanation of LBP, is the role of spinal stability. The stability of the spine depends on three subsystems: the spine, the muscles of the spine, and neural control. In CLBP subjects, spinal stability is assumed to be reduced due to loss of passive (vertebrae, intervertebral discs, ligaments) or active (muscles and tendons) control [26]. Therefore, changes in motor control are expected to compensate for this loss of stability [9, 54]. There is a variety of patterns of neural motor control in non-specific CLBP subjects, which suggest the existence of significantly different background mechanisms of pain [12] that may have a significant impact on clinical rehabilitation procedures. The characteristic of the lack of clinical stability for LBP subjects is that it affects each of the three subsystems and can be disrupted by any of them, resulting in reduced spine stiffness or inability to maintain balance in case of unexpected trunk displacements [74, 75]. One of the proposed mechanisms to compensate for spinal instability is to reduce the neutral zone (NZ) through muscle activation or gradual stiffening of the spine, such as osteophyte formation - ossification [26].

In the context of sEMG measurements, it is also important to mention that the stabilizing role of spinal muscles cannot be easily and directly measured given that they are deep muscles.

2.2.4 Motor adaptation to pain

The motor pain-adaptive model is based on four assumptions [7]:

1. muscle activity is redistributed within muscle and between muscles, unlike previous simplified muscle inhibition or excitation patterns
2. mechanical behavior is changed in a variable way in order to protect the tissue from further pain or injury (or danger of pain or injury)
3. changes occur at several levels of the motor system and can be complementary, cumulative, or opposite
4. benefits are shown in the short term, but with the possibility of developing long-term (negative) consequences due to factors such as increased load, reduced mobility, or reduced variability

2.2.4.1 Redistribution of muscle activity

In experimentally induced pain, a decrease in the frequency of activity triggering in motor units was observed, which was interpreted as a phenomenon of inhibition. However, recent work showed that the level of developed strength is maintained the same by employing additional motor units that were not active during contractions without induced pain [76]. This cannot be explained by uniform inhibitory mechanisms from the same set of motor neurons, but by

a redistribution of activity within the muscle involving new motor units despite the reduced frequency of activation of other (original) motor units [7].

In some body systems, especially those with significant redundancy such as trunk muscles, a spatial redistribution of muscle activity between muscles is observed. Thus, for example, delayed or reduced muscle activity of *m. transversus abdominis* is followed by the increase in activity of other abdominal and back muscles as part of postural adjustment before performing arm movements [27]. It should be noted that the total activity of the trunk muscles was increased, but it was achieved by different patterns of increase and decrease of muscle activity, which significantly depends on the individual subject.

2.2.4.2 Protection from pain or injury

Changes in mechanical behavior can be monitored during the current (acute) and chronic phases. During the acute phase, the motor system allows the nervous system to respond to and remove or reduce threatening painful stimuli (e.g., mechanical, chemical, thermal) toward the tissue. During the chronic phase, the motor response may be less meaningful, less accurate, or even unnecessary because the threat to the tissue may be less significant because of various physiological and psychological circumstances that alter the strength of the painful stimulus. Thus, the pain a person experiences does not necessarily correspond to the actual painful stimulus, and the pain does not necessarily reflect damage or potential tissue damage. Maintaining motor adaptation in chronic pain does not necessarily bring benefits to the body [7, 27].

The redistribution of activity among the trunk muscles also changes the kinematic and mechanical properties of the spine. The overall gross performance is preserved, in the sense that the movements are performed, but the quality of movement is reduced (increased spinal stiffness, reduced speed control, reduced ability to prepare the trunk for upcoming movements) which can have negative consequences for the individual [27].

2.2.4.3 Changes at multiple levels of the motor system

One of the changes or adjustments is the preparation (planning) of contractions (*anticipatory contractions*). A nice example is walking down the stairs where it is observed that the activity of the gluteal muscles precedes the contact of the feet. At the same time, the higher the step, the earlier and higher the activity will be. What is observed in subjects in whom foot contact with the step causes pain is that the motor-adaptive pain-protective mechanism will use a strategy normally used for higher-load tasks. Simply put, going down a 5 cm step will activate mechanisms as if it was a 15 cm step, through the mechanisms of preparation and planning of motor control [7]. Similar adjustments, that can be seen during arms movements, are present in order to avoid affecting the balance of the body (e.g. planned premature activity *m. erector spinae*) [77].

The level of these changes is shown to reach all the way to changes in the cerebral cortex in the form of cortical reorganization. The map of motor evoked potentials (MEP) in the cerebral cortex thus truly differs between LBP and healthy subjects [78].

It can be concluded that the overall effect on the outcome of the motor system depends on the contribution of individual components throughout the motor system and that it can be distinguished between individuals and tasks thus possibly explaining some variability in research results.

2.2.4.4 Short-term benefits, long-term consequences

These adjustments affect the quality of movement (increase spine stiffness and reduce the ability to control sudden movements), increase the load on tissues or change the load distribution (which can lead to tissue irritation), and reduce the variety of movements resulting in loading always the same structures during the same movements. One of the basic tenets of the new theory is that motor adaptation to pain is actually less important or even harmful in the long run. This is in line with current pain theory, which suggests that when pain exists and after the period during which the tissue has healed, the actual physical or mechanical causes of pain at the tissue level become less important and the mechanisms of maintaining or eliminating pain under central nervous system control become more important. Simply put, there is a difference between the actual level of “damage” to the tissue and our “feel” of pain in the brain. This is influenced by both physiological and psychological factors. Excessive pain is most often accompanied by a psychological factor of fear. Decreased pain sensation (although the "damage" to the tissue is significant) is often a long-term adjustment mechanism where the level of pain is modulated by the central nervous system [7].

2.3 Insights and ambiguities

In the study of LBP, different types of inputs, information, and features are examined, both electric (i.e. sEMG signals and myoelectric quantities), non-electric ones, or different metadata. Some of the commonly used not-electric parameters are time endurance and kinematic parameters (like a range of motion - ROM). In LBP studies it is very common for patients to fill out questionnaires in which they subjectively assess their health condition regarding LBP and the level of pain they feel. One such generally accepted and widespread questionnaire is the *Oswestry Low Back Disability Index* (ODI), which provides a statistically significant distinction between LBP and the control group [12, 40, 79]. Another one is *Roland-Morris Disability Questionnaire* (RDQ), which consists of 24 statements that the subjects should choose from or not, with respect to whether they correspond to their LBP condition. The total score ranges from 0 to 24, where 0 indicates the absence of incapacity caused by LBP. Another indicator

commonly employed is *Visual Analogue Scale* (VAS), used for pain scoring with the handwritten mark set on a 0 to 10 cm continuous line, where 0 denotes "no pain" and 10 "worst pain". All these measures are the relating to the degree of disability caused by LBP [54, 80].

It should be also noted that in the process of LBP categorization (classification) or diagnostics correlation, additional parameters related to the anthropometric and socio-demographic characteristics of the patients are used as well.

In this PhD thesis, the focus is put on acquiring and analyzing LBP characteristics by means of sEMG. Many years of sEMG signals research led to a better understanding of the corresponding properties in the temporal, frequency, and spatial domains and have yielded various features that can be utilized to characterize various electrophysiological changes in muscles and background changes in neuromotor control. Thus, these features can be generally grouped into:

- •Time domain (TD) features: mean absolute value (MAV), the sum of absolute EMG values (integrated EMG), variance, root mean square (RMS), the sum of absolute differences (waveform length - WL), number of zero crossings (ZC), the sum of signal slope changes (SSC), autoregressive coefficients, etc.
- •Frequency domain (FD) features: mean frequency spectral power frequency (MNF), median frequency of spectral power distribution *median frequency* , MDF)
- •Time-frequency domain (TFD) features: coefficients of different wavelet-based transformations WT (continuous - CWT, discrete - DWT, stationary - SWT, wavelet packet - WPT) and other *T-F* representations, such as STFT (short-time Fourier transform) or empirical mode decomposition (EMD) combined with Hilbert–Huang transform (HHT) enabling derived instantaneous frequency measures in T-F domain [81]
- •Spatial domain (SD) features: e.g. experimental variogram as a measure of the degree of spatial correlation [82]

More details and additional features can be found in the review papers [83, 84, 85, 86, 87, 88].

In the alignment of the given overviews, our research group has also contributed to the general understanding and applicability of sEMG in fatigue assessment [81, 83, 89] as well as tackling LBP subjects classification and differentiation as part of continuous efforts to better understand and support more successful rehabilitation treatments [22, 90].

Finally, looking into the studies related to LBP issues, the following characteristic groups of features tend to clearly prevail:

- •Frequency-domain features (FD) - regression line intersection (initial MDF) and regression line slope (MDF slope)
- •Muscle activity - root mean square (RMS) or sum of absolute EMG values (integrated EMG)
- •Time endurance (time-to-fatigue) and developed muscle strength

•Differences between left and right side

2.3.1 Frequency-domain features

Initial MDF has been shown to have no unequivocal evidence of its clear or overriding discriminatory power in distinguishing between LBP and non-LBP subjects. Namely, the initial MDF is said to be highly correlated with subcutaneous tissue thickness and is therefore not a good measure of muscle endurance [91]. Also, it is shown that the initial MDF relationship between the LBP and non-LBP subjects is ambiguous. Some papers report a statistically significant difference, where for LBP, the initial MDF values are constantly higher compared to non-LBP [17], while some papers state that there is no difference between the groups [40]. For MDF slope, there is also no unequivocal confirmation of a clearly discriminant feature power, but statistically more significant results are reported, suggesting faster muscle fatigue (higher MDF slope) in the lumbar region for LBP subjects compared to non-LBP ones [31, 54, 92]. Furthermore, for non-LBP subjects, it is observed that the muscles in the lower lumbar region (L5) fatigue faster than the ones in the upper lumbar region (L3 and T9) [17].

2.3.2 Features related to muscle activity

The second set of features is related to tracking muscle activation, either through the RMS or integrated EMG variants. Measures of muscle activity show more clearly a possibility of distinguishing between LBP and non-LBP groups. More concretely, it is almost unambiguously shown that muscle activity in LBP subjects is more pronounced. In that sense, the LLES area, in the LBP group, shows more pronounced activity compared to the ULES area, and with greater muscle co-contraction compared to the control group [9, 30, 31]. Also, in LBP subjects, increased muscle activity in the lumbar region was observed during walking [93] as well as during lifting movements [94]. Somewhat different results, similarly to the validation of the initial MDF parameter, Kankaanpää *et al.* [40] reported that the same level of muscle activity was observed in both LBP and non-LBP groups. A characteristic of this study was that all examined subjects were female.

2.3.3 Time endurance and developed strength

Similar to the parameters of muscle activity, the time endurance to fatigue is shown to be a good indicator of the presence of LBP. Most papers clearly and with statistical significance ($p < 0.05$) state that the time to fatigue is shorter in LBP subjects [9, 31, 40], with time shortening as the load is increasing, with resulting differences becoming even more significant [49]. An interesting result was reported in the work of Larivière *et al.* [51] where healthy subjects fatigued

faster than those with LBP. However, this result is most likely attributed to significantly lower loads taken by LBP subjects when compared to their maximum strength. Related to endurance, the level of strength developed, has also been shown to be a good indicator of the presence of LBP. Subjects with LBP most often show reduced trunk muscle strength compared to healthy subjects [95, 96], although some highly cited papers state that levels of developed strength are similar in both groups [31]. In addition, the same paper [31] states that LBP subjects often have a delay in performing movements or exercises that require higher levels of strength.

2.3.4 Differences between left and right side

Another characteristic, in the analysis of LBP, is the difference between the left and right sides of the paravertebral muscles of the lumbar region. Many papers state that in healthy subjects these differences do not exist [10, 52], or perhaps a more correct interpretation is that these differences exist, but are not so pronounced [31]. For LBP subjects, most papers state that there are differences between left and right [15, 97, 98], although not always significant in relation to the non-dominant side or even in relation to subjects without LBP [31, 91]. It is interesting to mention that even in healthy subjects who engage in a specific physical activity, where predominantly one side is used, such as rimen rowers (each rower has one oar), differences between dominant and non-dominant side would be detected [99].

2.4 sEMG signal properties

One set of factors, which determine which signal processing approach to apply, depends on the mathematical and statistical properties of the signal itself, in this case, the sEMG signal. Thus, certain signal processing techniques have limitations when we talk about the possibility of analyzing the signal properties characterized by nonstationarity or nonlinearity.

The stochastic nature of sEMG signal can be nicely observed when compared to other bioelectric signals such as electrocardiography signal (ECG), thus giving a more intuitive notion of certain challenges posed in sEMG signal processing and analysis (Fig. 2.4).

2.4.1 Nonstationarity

Namely, taking into account that there are no satisfactory models that describe the deterministic nature of complex neural control of muscle contractions, the sEMG signal is most often considered a stochastic process [100]. Temporal changes in the amplitude and frequency distributions of EMG signals, and related statistical properties, violate the assumption of stationarity, which is one of the prerequisites for the applicability of Fourier transform (FT) based analysis techniques. When examining the properties of sEMG signals at smaller time segments (splitting

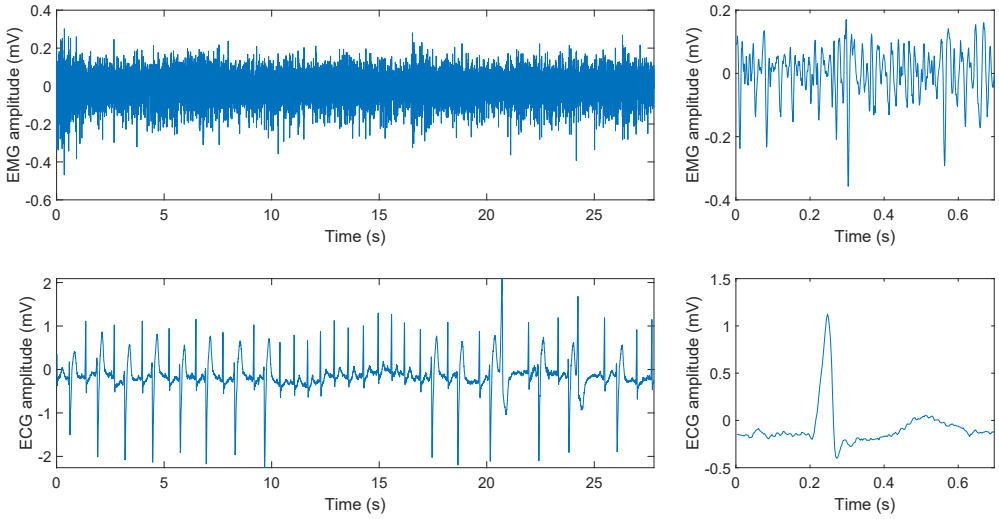


Figure 2.4: Comparison of EMG (above) and ECG (below) recordings of the same duration for recordings on longer (left) and shorter (right) segments.

signals into such shorter sequences), it is noticed that the statistical properties of EMG signals change more slowly and that conditions for wide-sense stationarity (WSS) are satisfied. These assumptions are primarily confirmed for myoelectric signals recorded during isometric contractions, which allows the application of the Fourier transform on short time segments (STFT) with a duration of $0.5 - 2$ s [101, 102]. However, if using the STFT method (2.1), time-frequency resolution constraints are introduced, with short EMG segments providing good time but poor frequency resolution, which is a consequence of the well-known *uncertainty principle* [103].

$$X(\tau, f) = \int_{-\infty}^{+\infty} x(t) g^*(t-\tau) e^{-j2\pi f t} dt \tag{2.1}$$

Dynamic contractions show that the assumptions of wide-sense stationarity are usually not met, and arise from the fact that during such contractions there is a change in the angle between the joints, changes in muscle fibers length, muscle strength, and speed, as well as changes in relative electrode position in respect to active muscle fibers [102, 104].

2.4.2 Nonlinearity

In addition to the nonstationary properties, the sEMG signal is generally characterized by nonlinearity properties as well [105, 106]. Moreover, Nieminen and Takala [107] suggest that myoelectric signals could be even better modeled as a result of nonlinear dynamical systems than as a result of random stochastic processes. Accordingly, various nonlinear parameters or techniques are proposed throughout the research papers that should enable the analysis of such

nonlinear properties of myoelectric signals. Thus, for example, it is proposed to use entropy as a nonlinear measure of signal complexity or the analysis in the form of fractal analysis and *recurrence quantification analysis* (RQA), as often cited in a number of review papers [83, 86]. Also, the assessment of muscle fatigue by utilizing the constructed muscle fatigue index (*fatigue mapping index*) by combining different sEMG parameters through the learning process using neural networks (using a multi-layer perceptron, MLP) [108] was also proposed, where it was shown that such proposed method does not show statistically significantly better results compared to linear methods [86]. As additional techniques for nonlinearity analysis, higher-order statistics (HOS) methods can also be mentioned, that allow measuring the deviation from linear and stationary properties in the signal, as well as the deviation from the Gaussian distribution [109]. It is worth mentioning the EMD, first introduced by Huang *et al.* [110], as a method that allows the analysis of nonlinear and nonstationary signals in the time domain, and is characterized by its adaptation to data (i.e. *data-driven, data-adaptive*). In the meantime, EMD as a technique has been very well accepted with applications in various engineering fields [111], and over time further improved [112, 113] with finding applications in different fields of data and signal processing and analysis, including sEMG signals [81, 84, 114].

2.5 Motivation and hypothesis

Research in this PhD thesis has been motivated by still an obvious lack of success in defining subgroups of LBP patients, consequently leading to the inability to define effective treatments for non-specific LBP patients [12]. Moreover, as stated in [25], it is considered the "Holy Grail" of low back pain research to establish such methods that would enable subgrouping of LBP patients, backed with a consistent medical interpretation. By examining the so far reported results and state-of-the-art, in this PhD thesis it is hypothesized that an additional contribution to successful subgrouping of LBP patients can be supported by introducing more comprehensive multifactorial LBP classification and clustering models, thus also capturing more individually-biased characteristics of LBP expressions, and avoiding a common pitfall of "one-size fits all" approach [8, 25]. Subsequently, another key ingredient in this proposed method is an attempt to reflect the LBP specificities directly within classification features themselves, more precisely, to model such features that (1) contain contextual information pertained to LBP, and (2) that are interpretable by medical experts. Namely, currently, it is still a predominant case that feature inputs into the classifiers are simplistic features or measures (e.g. straightforward RMS, MDF or rectified EMG values as utilized in [17, 22, 23, 43, 44]), meaning, not specific to LBP phenomenon itself and its complex coordination, co-activation and compensation mechanisms in any directly modeled manner. Thus, the overall set hypothesis in this thesis is that LBP complexity requires a more elaborated classification modeling approach to enable insightful in-

ference. This idea follows the concerns and insights established throughout the review studies by Dieën *et al.* [9]), Geisser *et al.* [41], and pace-setting works by Hodges [7, 27]. In this PhD thesis, the complexity focus is put on features modeling, not on the complexity of the acquisition systems or exercise protocols.

In order to validate the hypothesis and the proposed method construction, the following setup was opted: sEMG features only (thus avoiding bias from other non-myoelectric values), simple isometric trunk extension exercise (thus avoiding bias from dynamic functional tasks that have demonstrated certain LBP discriminating power), only four sEMG channels for myoelectric activity acquisition (thus avoiding contribution to discrimination success usually pertained to multi-electrode setups), classifiers and clustering methods employed *as-is* out-of-the-box (thus avoiding bias from additional methods optimization or fine-tuning). The rationale behind a such proposed setup is as follows: if this minimum setup can demonstrate success in detecting LBP patients and provide insights into LBP groups at hand (with further CLBP subgrouping) with satisfactory differentiation and clustering accuracy, alongside preserving the possibility to interpret the results, then it is reasonable to assume that the proposed approach and algorithmic procedure behind can withstand more general challenges with further enhancements and optimizations and even further contribute to the overall goal of enabling effective and meaningful subgrouping of LBP patients.

Finally, the aim of this PhD thesis is in providing a methodological background for building the clinical decision support systems (CDSS) that may be used by medical professionals in clinical practice as part of more concrete medical guided-based recommendation systems, with strong interpretable models. Moreover, such explainable and interpretable models are mandated by European Union's General Data Protection Regulation (GDPR), as pointed out in work by Jović *et al.* [115].

Chapter 3

SEMG recording and features extraction

In this chapter, as the first step in enabling this research, the acquisition of sEMG signal recordings from both the control (healthy) group and LBP patients is presented. Three different male cohorts of subjects were defined, based on the medical professional, resulting in three groups: healthy subjects (HS), subjects with non-specific chronic LBP (CLPB), and subjects with radiculopathy (RLBP) with thirty-seven, twenty-nine, and twenty-five subject respectively.

LBP was induced by performing an isometric trunk extension exercise in the Roman chair configuration by a clearly defined measurement protocol followed by the respective subjects and supported by medical staff. The respective exercise and sEMG recording were stopped when the subject was no longer able to sustain the trunk extension position or a time limit of 180 s was reached. sEMG recordings on two bilateral sites (four channels) in the lumbar region were acquired through the electromyograph system with wireless surface EMG probes. This resulted in myoelectric activity recorded for upper lumbar erector spinae (ULES) and lower lumbar erector spinae (LLES) muscle sites, for L1-L2 and L4-L5 vertebrae positions, respectively.

Subsequently, the acquired sEMG recordings, sampled at 1000 Hz, were then further mathematically processed employing signal processing techniques, where the following initial fourteen (simple) raw sEMG features were calculated and analyzed: zero crossing (ZC), signal slope change (SSC), Willison amplitude (WAMP), mean absolute value (MAV), integrated EMG (IEMG), variance (VAR), root mean square (RMS), waveform length (WL), log detector (LD), kurtosis (KURT), skewness (SKEW), permutation entropy (PE), power spectrum median frequency (MDF), and relative variance difference (RVD).

3.1 Materials and Methods

3.1.1 Subjects

Three different cohorts of male volunteers were involved in this research based on the presence of low back pain or related pathologies. All subjects were examined individually by the medical professional and, following up their medical history track and diagnostic check, each subject was assigned to one of the three groups: HS, non-specific CLBP, and RLBP. Inclusion criteria for CLBP were defined as daily or almost daily pain that lasted at least three months prior to measurements [12, 30, 50]. For RLBP patients, the main inclusion criteria were clinical symptoms of radiculopathy with a positive Lasegue sign on one side lasting at least fourteen days [18, 57, 58]. The overall exclusion criteria for subjects were spinal deformation, spinal surgery, spondylolisthesis, spinal stenosis, or spinal injuries, altogether with the indication that there were no accompanying systemic diseases [18, 30]. Following the defined procedure, ninety-one male subjects were detected as suitable for measurement protocol and sEMG data acquisition. Alongside collecting the biometric and demographic details, subjects were provided intake of standard questionnaires, namely, Oswestry Low Back Pain Disability Questionnaire (ODQ), Roland-Morris Disability Questionnaire (RDQ), and Visual Analogue Scale (VAS) for pain scoring, to assess own LBP related functional disability and pain perception status [12, 30, 40, 54]. On top of these measures and scores, the Lasegue test (FFK), also known as the straight leg raise (SLR) test was measured [116]. FFK is a fundamental maneuver during the physical examination of a patient with lower back pain. It aims to assess for lumbosacral nerve root irritation [117]. All these metadata are summarized in Table 3.1. The data has been collected at the Biomechanical laboratory of Polyclinic for Physical Therapy and Rehabilitation, Pula, Croatia. The collection time was in the morning from 9:00 am to 12:00 pm.

The whole experiment was approved by the Ethics Committee of the University of Zagreb, Faculty of Electrical Engineering and Computing, and informed consent was received from each subject.

3.1.2 Measurement Protocol and sEMG recording

Measurement protocol was based on the isometric trunk extension exercise in the configuration of Roman chair (a variant of *Biering-Sørensen* test) [22, 45, 54], as shown in Fig. 3.1. Myoelectric muscle activity was recorded with FREEEMG (BTS, Milano, Italy) system with wireless surface EMG probes enabling free movement of subjects during the measurement. Each surface EMG probe applied had a pair of pre-gelled Ag-AgCl 10 mm diameter electrodes (Ambu-Blue, Sensor, and Ballerup, Denmark). Prior to applying the electrodes onto the skin, the respective

Table 3.1: Number of subjects (*N*) for HS, CLBP, RLBP cohorts with demographic (age) and biometric (height, weight, body mass index - BMI) details, together with Oswestry Disability Index (ODI), Roland-Morris Disability Questionnaire (RDQ) outcome and Visual Analogue Scale (VAS) pain scoring, Lasegue test (FFK), for left (sin) and right (dex) side, and $T_{Exercise}$ as a measured trunk extension endurance time, are shown as the mean value and standard deviation (mean $\pm$ std).

	HS (N = 37)	CLBP (N = 29)	RLBP (N = 25)
Age (years)	32.62 $\pm$ 9.17	34.1 $\pm$ 10.12	36.4 $\pm$ 9.0
Height (cm)	180.76 $\pm$ 5.18	184.03 $\pm$ 7.63	182.56 $\pm$ 5.52
Weight (kg)	82.84 $\pm$ 8.42	87.24 $\pm$ 10.39	86 $\pm$ 9.21
BMI (kg/m^2)	25.35 $\pm$ 2.23	25.70 $\pm$ 2.0	25.81 $\pm$ 2.64
ODI (0-100)	7.14 $\pm$ 8.1	15.59 $\pm$ 9.49	33.04 $\pm$ 14.1
RDQ (0-24)	1.86 $\pm$ 2.57	5.34 $\pm$ 3.91	11.24 $\pm$ 4.76
VAS (0-10)	1.65 $\pm$ 2.02	2.74 $\pm$ 1.84	5.6 $\pm$ 2.14
FFK sin (°)	80.59 $\pm$ 8.35	74.90 $\pm$ 10.71	51.24 $\pm$ 16.37
FFK dex (°)	80.43 $\pm$ 8.01	74.48 $\pm$ 11.29	55.44 $\pm$ 21.75
$T_{Exercise}$ (s)	139.95 $\pm$ 43.74	105.50 $\pm$ 48.38	64.72 $\pm$ 39.66

surface was prepared in accordance with SENIAM recommendations with surface cleansing and electrode-skin impedance controlled (<5 kΩ) [118]. The main characteristics of the EMG recording system were given by: differential amplifying with bandpass filtering (20-400 Hz), differential input impedance >100 MΩ, CMRR >100 dB (at 65 Hz) with 1000 Hz sampling frequency using a 12-bit A/D converter. Outside of the 20-400 Hz frequency range, the contribution to the sEMG spectrum is deemed negligible [119].

To commence a testing procedure, each subject was familiarized with the procedure, a tilting device, and instrumentation. Further, each subject was asked to stand upright until electrodes were positioned. Four sEMG channels were acquired in the paraspinal lumbar region, placed over the upper lumbar erector spinae (ULES) and lower lumbar erector spinae (LLES) muscle sites, bilaterally (left and right) in respect to the vertical spine axis (Fig. 3.2). While standing in the upright position without footwear, the distance between the floor and *anterior superior iliac spine* was measured. Depending on the measured distance, the standing pad of the tilting device was adjusted so that the toes, back of the lower leg (above the Achilles tendon), and pelvis (together with the upper thigh) became the only body parts in contact with the tilting device and thus creating the supporting points. The subject was instructed to stand on the tilting device and to hold hands crossed having palms placed on his chest. Upon the subject's verbal confirmation, the medical staff gradually tilted the device until the horizontal position was reached. To ensure static contractions of lower back muscles each subject was asked to maintain in a horizontal position as stable as possible. The weight of the subject's upper body was used to induce muscle fatigue. Subjects were asked to sustain the isometric contraction as long as possible or until a maximum exercise duration of 180 seconds was reached. When

exercise end was reached, medical staff returned the subject to an upright position [22].

Figure 3.1: A tilting device in the Roman chair configuration was used for testing with the subject placed in a trunk extension position.

3.2 Raw Features

3.2.1 SEMG signal preparation

Surface myoelectric signals are generally characterized by non-stationary [81, 120, 121] and non-linear [105, 106, 107] stochastic properties as a result of not having satisfactory models describing a deterministic nature of complex neural control for muscle contractions [100]. Thus, different segments within sEMG time sequence have different statistical properties and carry different pieces of information relevant for respective detection and inference of LBP conditions. This can be well noticed in multi-electrode sEMG topography measurements with dynamic exercise tasks [52]. Therefore, it would be plausible to analyze consecutive portions of sEMG signals to infer certain localized time-dependent properties that would otherwise be lost. Alongside, within shorter time chunks of sEMG signal, statistical properties are changing much slower and wide-sense stationarity (WSS) can be assumed. This primarily holds for myoelectric signals derived from isometric contractions with duration of $500-2000$ ms [101, 102, 120], whereas for dynamic type of contractions these assumptions may be violated [101, 104, 122].

Based on these insights, in this research time chunks of $L = 1000$ ms were applied to calculate primary features with overlapping step $s = 50$ ms (the equivalent of 95% overlapping between neighboring time chunks).

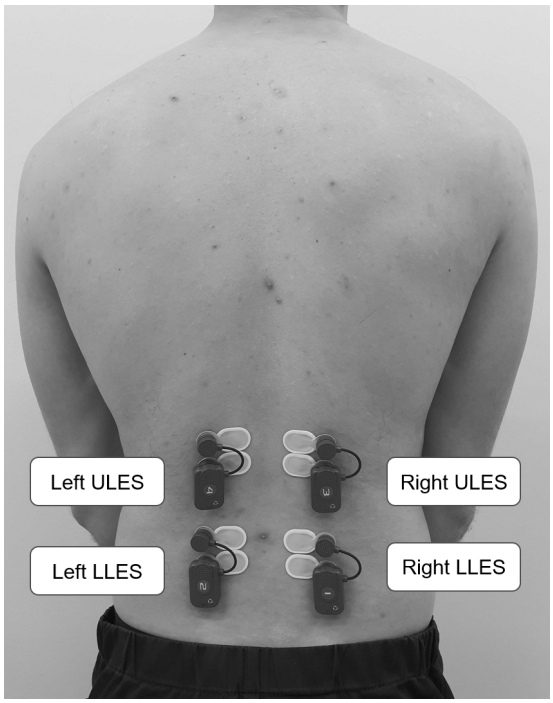


Figure 3.2: Placement of two bilateral pairs of wireless electrodes on the L1-L2 (ULES) and L4-L5 (LLES) positions. Altogether four channels.

3.2.2 Primary Features Extraction

As the first step in the proposed procedure, simple raw features (also referred to as the *primary features* further in the text) were calculated directly from the raw sEMG signals. These primary features in subsequent steps were used to create more LBP-meaningful secondary features. Depending on the domain where features are calculated, they are generally categorized as time domain (TD), frequency domain (FD), time-frequency domain (TFD), and spatial domain (SD) based features, as presented in section 2.3

In this PhD thesis, time domain-based features were mostly opted for due to their simplicity and speed of calculation. The full list of primary features being utilized, with the originating domain and types of information carried by the respective feature, is given in Table 3.2. Thirteen TD features: ZC, SSC, WAMP, MAV, IEMG, VAR, RMS, WL, LD, KURT, SKEW, PE, and RVD were utilized. The only FD feature employed was MDF. This feature list, among others, contains all features presented by Hudgin [123] or Du [124]. Although respective feature vectors (namely, Hudgin's and Du's) are primarily used in real-time applications exploiting sEMG, like human-machine interfaces (HMI), it was worth also exploring and comparing the applicability of similar feature sets in the domain of LBP detection and patients differentiation.

Detailed overviews of these and other common features created in TD, FD, TFD, and SD can be found in [85, 86, 87, 88]. The presented features, irrespective of being created in time or (time-)frequency domain, can reflect certain properties that are natively expressed in another domain. For instance, some of the TD features, like ZC or SSC are closely related to the

Table 3.2: Primary features used in this research with feature name, feature domain (TD, FD), and type of information carried (frequency, energy, force, activation, complexity, non-linearity).

ID	Domain	Feature name	Type of information
1	TD	Zero crossing (ZC)	Frequency
2	TD	Signal slope change (SSC)	Frequency
3	TD	Willison amplitude (WAMP)	Frequency, activation, force
4	TD	Mean absolute value (MAV)	Energy, force
5	TD	Integrated EMG (IEMG)	Energy, force
6	TD	Variance (VAR)	Energy, force
7	TD	Root mean square (RMS)	Energy, force
8	TD	Waveform length (WL)	Energy, complexity
9	TD	Log detector (LD)	Energy, force
10	TD	Kurtosis (KURT)	HOS (non-linearity)
11	TD	Skewness (SKEW)	HOS (non-linearity)
12	TD	Permutation entropy (PE)	Non-linearity, complexity
13	FD	Power spectrum median frequency (MDF)[a]	Frequency
14	TD	Relative variance difference (RVD)[b]	Activation, complexity

[a]MDF is the only FD feature used in this research.

[b]RVD is a newly introduced feature in this research.

frequency properties of the respective signal, that are originally observed in FD. Also, as part of this PhD thesis, an additionally constructed RVD feature was introduced and is presented in more detail in section 3.2.2.14, together with a less commonly used PE feature (section 3.2.2.12).

Our primary features are calculated on a sequence of raw sEMG time chunks ($L = 1000$ ms, $s = 50$ ms) where for each feature (Table 3.2) the output is a new sequence, effectively sampled due to step s used. This new feature sequence then holds more pronounced respective domain (TD or FD) information, by reflecting the progression of myoelectric properties that are expressed in terms of the specific primary feature used.

3.2.2.1 Zero crossing (ZC)

Number of zero crossings is the measure of the frequency information of an EMG signal defined in the time domain. The general formula may also include a threshold criterion to exclude the

effects of small-amplitude or noise fluctuations.

$$ZC = \sum_{i=1}^{N-1} f\left(-x_i \times x_{i+1}\right)$$
$$f\left(x\right) = \begin{cases} 1, \text{ if } x \geq 0 \\ 0, \text{ otherwise} \end{cases} \tag{3.1}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.2 Signal slope change (SSC)

The number of signal slope changes is another method that represents the frequency information of the EMG signal. Three consecutive time samples (three values of the sEMG signal) are needed for the calculation, and the mentioned mathematical formulation can be observed in the context of the sum of the maximum and minimum counters of a given signal.

$$SSC = \sum_{i=2}^{N-1} f\left[\left(x_i - x_{i-1}\right) \times \left(x_i - x_{i+1}\right)\right]$$
$$f\left(x\right) = \begin{cases} 1, \text{ if } x \geq \text{threshold} \\ 0, \text{ otherwise} \end{cases} \tag{3.2}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.3 Willison amplitude (WAMP)

Willison amplitude is also a measure of the frequency content of an EMG signal, defined in a similar way to ZC and SSC. The measure represents the number of differences in the amplitudes of two adjacent segments of the EMG signal that exceed a predefined threshold. Also, this measure is associated with the phenomenon of motor unit action potential triggers as well as with the force of muscle contraction.

$$WAMP = \sum_{i=1}^{N-1} f\left(\left|x_i - x_{i-1}\right|\right)$$
$$f\left(x\right) = \begin{cases} 1, \text{ if } x \geq \text{threshold} \\ 0, \text{ otherwise} \end{cases} \tag{3.3}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.4 Mean absolute value (MAV)

Mean absolute value is one of the most commonly utilized measure in the analysis of EMG signals, and is especially applied onto sEMG signals to control prosthetic limbs. Also to note, many similar values, utilized in EMG analysis, were derived from the basic MAV formulation.

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{3.4}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.5 Integrated EMG (IEMG)

Integrated EMG is a common measure employed to detect onset of muscle activity (contraction) in clinical applications. It is very similarly defined as MAV and is calculated as the sum of the absolute values of the amplitudes over the observed section of (EMG) signal with duration N.

$$IEMG = \sum_{i=1}^{N}|x_i| \tag{3.5}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.6 Variance (VAR)

Variance is yet another measure of strength or energy. In general, variance is calculated as the mean square deviation of an array, from its mean. However, since the mean value of the sEMG signal is very close to zero (at the level of precision computer error) 10^{-10}, the mathematical expression can be simplified:

$$VAR = \frac{1}{N-1}\sum_{i=1}^{N}x_i^2 \tag{3.6}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.7 Root mean square (RMS)

Root mean square is another very popular feature in EMG signal analysis. The mathematical formulation is very similar to the expression for variance, i.e. practically identical to the expression for standard deviation (with the previously mentioned assumption of the mean value of the EMG signal close to zero). RMS well reflects conditions related to contractions with constant force and without the development of fatigue. Although, it should be noted, the occurrence of fatigue is characterized by an increase in EMG amplitude, and consequently RMS values on a

given segment of EMG signals.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{3.7}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.8 Waveform length (WL)

Wavelength is a measure of the complexity of an EMG signal. It is defined as the sum of the EMG wavelengths over the time segment.

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{3.8}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.9 Log detector (LD)

Log detector is used as one of the estimators for muscle contraction force and, as the name suggests, is based on a logarithmic function.

$$LD = e^{\frac{1}{N}\sum_{i=1}^{N} \log(|x_i|)} \tag{3.9}$$

where x stands for the sEMG time sequence of length $N = 1000$ ($L = 1000$ ms).

3.2.2.10 Kurtosis (KURT)

Kurtosis is the higher order statistics (HOS) feature that gives a measure of the deviation of a given distribution (random variable) from the normal Gaussian distribution (note: kurtosis value for Gaussian distribution is 3). Related to this definition, this feature inherently also provides an insight into the, so-called, *peakedness* in terms of detecting *sub-Gaussian* distributions ($KURT < 3$, e.g. multimodal distributions) or *super-Gaussian* distributions ($KURT > 3$, e.g. as for *Laplace's* distributions).

Kurtosis is calculated as the fourth standardized moment (the measure of the expectation, $E[]$, of the fourth-order cumulant divided by the square of the second-order cumulant), and given the assumption of the zero-mean value of the EMG signal, the expression is reduced to:

$$KURT = E\left[\left(\frac{X - \mu_x}{\sigma}\right)^4\right] \approx \frac{E\left[X^4\right]}{E\left[X^2\right]^2} \tag{3.10}$$

where higher order statistics is calculated on the sEMG time sequence x of length $N = 1000$.

3.2.2.11 Skewness (SKEW)

Skewness is a feature of higher order statistics, a measure of third order cumulant. This measure can also be seen as a measure that provides insight into the asymmetry of a given distribution. Where "negative skew" has a tail on the left side of the distribution (*left-tailed*), and "positive slant" has a tail on the right side of the distribution (*right-tailed*). Given the assumptions of the zero-mean value of the EMG signal, the expression for the asymmetry coefficient is reduced according to:

$$SKEW = E\left[\left(\frac{X-\mu_x}{\sigma}\right)^3\right] \approx \frac{E\left[X^3\right]}{E\left[X^2\right]^{3/2}} \tag{3.11}$$

where higher order statistics is calculated on the sEMG time sequence x of length $N = 1000$.

3.2.2.12 Permutation entropy (PE)

Entropy is a non-linear measure of signal complexity. There are different formulations and parameters that can be used to estimate the entropy for the given signals. Permutation entropy was chosen - PE [125]. With PE the relative occurrences of different patterns or *motifs* can be quantified. The strong side characterizing PE is the ability to tackle non-linear properties of sEMG signals together with demonstrating robustness in eliminating potential negative effects coming from interferences. The main downside is the computational complexity with calculation taking roughly two orders of magnitude more time compared to other TD features [125, 126].

The core of PE is in choosing n consecutive points of samples and constructing a series of n-dimensional sequences (by shifting in time). Alongside, a set of potential sequence patterns of length n (set size of magnitude $n!$) is constructed. The n-dimensional sequence is said to fall into a specific pattern if it exhibits the same amplitude's rise or fail trend in time. Subsequently, the probability statistics for the permutations and combinations over the entire time series is calculated [126]. The probability is represented as $p(\pi)$, where π stands for different permutations. The expression is given with the following,

$$PE = -\sum_{\pi=1}^{n!} p(\pi)\ln(p(\pi)). \tag{3.12}$$

The given formulation and computational implementation is following the original work done by Bandt and Pompe [127]. In order to calculate PE, it is necessary to select n (n-th order of permutation entropy) where the higher order generally enables better estimation of the complexity of the underlying dynamic system. Though, on finite size signals, too high order can result in underestimating the system complexity because many of the patterns (part of the overall permutation set) are likely not showing up on the finite size time chunk. It is shown that

an empirical formula $5 \times (\mathrm{n}+1)! < L$ for calculating order n provides good results [128]. Thus, in this case, with the time chunk length of $L = 1000$ ms, $n = 4$ was chosen.

To additionally supplement, the computational time required to calculate the PE feature is 2 orders of magnitude larger than for other features. As part of this research, the *Fuzzy entropy* nonlinear feature was also considered, but its further use was abandoned as preliminary results did not show a significant contribution to the success of the classification. In addition, this feature also required an increased computational time.

3.2.2.13 Power spectrum median frequency (MDF)

The median frequency is defined as the frequency at which the power spectral density (PSD) is divided into two parts of equal energy. MDF has been shown to be closely related to the biochemical and physiological aspects of muscle fibers conduction velocity (CV) and to reflect well the biological changes in muscle tissue associated with muscle fatigue. Also, MDF is shown to be related to the (histological) composition of muscle fibers [50]

$$\int_{f1}^{MDF} P(f)\,df = \int_{MDF}^{f2} P(f)\,df \tag{3.13}$$

where $f1$ and $f2$ stand for the highpass and lowpass cut-off frequencies defining the bandpass range. In this setup, $f1$ and $f2$ correspond to 20 HZ and 400 Hz, respectively.

3.2.2.14 Relative variance difference (RVD)

Relative variance difference is an additional feature constructed as part of this PhD thesis and first-time introduced as part of the respective research. It is constructed with a goal to enable detecting and emphasizing the relative changes in localized sEMG signal energy (related also to amplitude changes) and specific to localized contraction properties. The relative measure is introduced having in mind the character of myoelectric manifestations recorded on the surface of the skin where the absolute values of sEMG signals may quite differ not only among different subjects but also among different channels recorded from the same subject. The presented measure (feature) is given by

$$\begin{aligned} RVD(x_k) &= \frac{Var(x_{k+}) - Var(x_{k-})}{Var} \\ &= \frac{\frac{1}{W-1}\left(\sum_{i=k}^{k+W} x_i^2 - \sum_{i=k-W}^{k} x_i^2\right)}{\frac{1}{N-1}\sum_{i=1}^{N} x_i^2} \end{aligned} \tag{3.14}$$

where for k-th step in the time sequence, a variance difference around time sequence point x_k is calculated for the prior ($Var(x_{k-})$) and posterior ($Var(x_{k+})$) variance of time segment of length W. In this research, segment length $W = 1000$ ms was used. Variance difference is normalized against the overall sEMG signal (channel) variance.

Chapter 4

Contextual features modeling

The rationale behind introducing additional feature construction, on top of the primary features, was the intention to provide functional and descriptive features that reflect the concrete LBP domain-related challenges with the ability to interpret the results. One of the main tasks was to enable detection of LBP patients and consequently provide more specific differentiation of patients within LBP groups as, nowadays, this poses one of the main challenges in dealing with LBP patients and rehabilitation treatments [8, 9, 12, 25]. To achieve this, primary features were used to construct contextual features (also referred to as the *secondary features* further in the text) that are organized into the following feature groups:

1.Coordination measures
2.Co-activation triggers
3.Trends
4.Fatigue-related indices

4.1 Coordination measures

Coordination-related features were introduced and proposed to detect a relationship among the channels in different muscle locations. Previous studies reported differences or asymmetry in myoelectric patterns between left and right for patients with LBP compared to healthy subjects [16, 52, 97]. The main coordination relationship investigated in this research was the bilateral coordination, i.e. time-related dependency between the left and right side of paraspinal muscles (ULES and LLES) in the lumbar region (Fig. 3.2). In that course, relationships for the left-right ULES, as well as left-right LLES, were tracked separately, resulting in two features per subject per single primary feature employed.

The aforementioned bilateral muscle relationships were established by employing different distance metrics to measure how closely each left side time-sequence is in alignment with the same corresponding sequence (i.e., derived from the same primary feature) on the right side.

This way, a macroscopic muscle coordination indicator of myoelectric similarity between muscle locations (channels) was defined. Calculations were conducted only between the muscle location pairs for the same subject, thus sequences were inherently of the same length. Following distance metrics were employed:

1.Euclidean distance

2.Correlation distance

3.Dynamic time warping (DTW)

4.Spearman distance

5.Mutual information (MI)

Thus, overall ten secondary features per single primary feature were constructed, $N_{Co} = 10$.

4.1.1 Euclidean distance

Euclidean distance is a classic, very common distance measure between two points in the vector space. In this case, distance measure was directly employed onto two, same length, primary features time sequences, as in

$$d_{Eucl}(p,q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \tag{4.1}$$

where p and q stand for the respective feature sequences.

Empirical results have demonstrated Euclidean distance to provide very good results in many applications. Moreover, considering the trade-off between the calculation speed and the accuracy, this measure has often proved to be the preferred choice. In this case, the Euclidean distance measure is applied on two hyper-dimensional vectors, thus posing some challenges in terms of applicability [129].

The calculated distance measure value (4.1) is subsequently used as one of the features as part of the presented classification and differentiation process, thus effectively used for comparison among subjects in the data set. For that reason, to enable meaningful comparison among subjects, the impact of different time sequence lengths should be eliminated by introducing a normalization or correction factor. Given the example:

$$p1 = [1, 1, 1, 1],\ q1 = [1, 5, 1, 5],\ d_{Eucl}(p1,q1) = 5.657$$
$$p2 = [1, 1, 1, 1, 1, 1],\ q2 = [1, 5, 1, 5, 1, 5],\ d_{Eucl}(p2,q2) = 6.982$$

we want distance measures for both sequence pairs to be the same, $\hat{d}_{Eucl}(p1,q1) = \hat{d}_{Eucl}(p2,q2)$. This is achieved with the following normalization expression:

$$\hat{d}_{Eucl}(p,q) = \sqrt{\frac{\sum_{i=1}^{n} (p_i - q_i)^2}{n}} \tag{4.2}$$

where now the calculated distance values, $\hat{d}_{Eucl}(p1,q1) = \hat{d}_{Eucl}(p2,q2) = 2.828$, are the same.

4.1.2 Correlation distance

Correlation distance is calculated according to the formulation

$$d_{Corr}(p,q) = 1 - \rho(p,q). \tag{4.3}$$

whereas the expression in (4.3) is derived from the Pearson correlation coefficient given by

$$\rho(p,q) = \frac{\sum_{i=1}^{n}(p_i - \bar{p})(q_i - \bar{q})}{\sqrt{\sum_{i=1}^{n}(p_i - \bar{p})^2 (q_i - \bar{q})^2}} \tag{4.4}$$

where p and q stand for the respective primary feature sequences and their mean values, $\bar{p}$ and $\bar{q}$, respectively. The correlation distance measure as such does not require any additional normalization for comparison across subjects.

4.1.3 Dynamic time warping

DTW is a technique that enables non-linear mapping of one signal onto another one by finding the optimal path W_0 that minimizes the warping cost [130], as in

$$d_{DTW}(p,q) = \min_{w} \left\{ \sqrt{\sum_{k=1}^{K} w_k} \middle/ K. \right. \tag{4.5}$$

A warping path W is a contiguous set of matrix elements ($w_k = (i,j)_k$) that characterizes a mapping between our two primary feature sequences p and q, resulting with the minimal root sum of Euclidean distances between individual sequence points, namely $d(p_i,q_i) = (p_i - q_i)^2$ (Fig. 4.1). Alongside the warping path, of length K, being calculated, the output of this procedure is also the similarity measure between the two sequences $d_{DTW}(p,q)$. The problem of finding the respective optimal warping path is solved by utilizing dynamic programming techniques.

A trivial example of DTW distance measure and similarity detection can be illustrated with the following two sequences (Fig. 4.2)

$$p = [0, 0, 5, 0, 0, 0, 0, 0]$$
$$q = [0, 0, 0, 0, 0, 5, 0, 0]$$

where for DTW a measured distance between two sequences is $d_{DTW}(p,q) = 0$, thus perfect alignment. In the case of Euclidean distance, it results with $d_{Eucl}(p,q) = 7.07$. In a core, this

demonstrates the basic difference and advantage of DTW compared to a straightforward Euclidean metric, but at the cost of higher computational complexity two to three times orders of magnitude). This especially can become an important constraint when dealing with lengthy sequences [131]. This, in fact, has been DTW's main disadvantage for more common usage, although this technique has been well known for a long time in the field of speech recognition [132] and applications in many other fields where analysis of (time) sequences is required, e.g. gait analysis [133] or matching among different samples is required (e.g. shape matching, signature recognition) [134]. Nowadays, more powerful computational systems and advancements in DTW-based algorithms have opened a new perspective for employing this powerful technique in many other fields and applications [131].

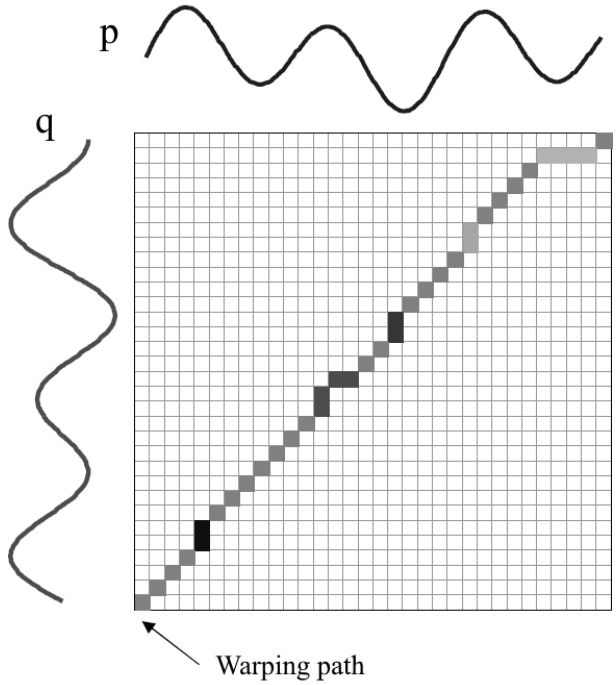


Figure 4.1: DTW warping path for two sequences with a contiguous set of matrix elements. The material is taken from [135] and adjusted. Available online `http://didawiki.cli.di.unipi.it/lib/exe/fetch.php/dm/time_series_from_keogh_tutorial.pdf`

4.1.4 Spearman distance

Spearman distance is calculated as Spearman correlation which corresponds to Pearson correlation between the rank values of those two variables (in our case two feature sequences) [136]. Pearson correlation primarily deals with linear relationships successfully, whereas Spearman correlation assesses the monotonic relationships (which can be non-linear). That said, Spearman correlation results in maximum values (+1 or -1) for cases when sequences with no repeating values are acting as a perfect monotone function against each other, even if not being in a

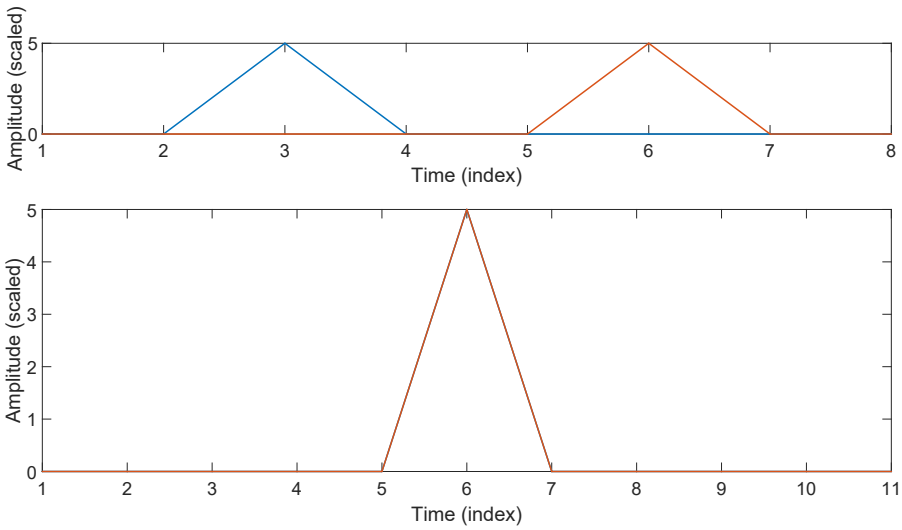


Figure 4.2: Two simple sequences left (blue) and right (orange) are submitted to DTW (top). DTW warping results in perfect alignment and pattern match (bottom) with distance $d_{DTW}(p,q) = 0$, compared to Euclidean distance of $d_{Eucl}(p,q) = 7.07$.

linear relationship [137].

Spearman distance uses rank values for calculation (i.e. values that correspond to the relative position of an individual point within the sequence after sorting):

$$d_{Spear}(p,q) = 1 - \frac{\sum_{i=1}^{n}(r_{pi} - \bar{r}_p)(r_{qi} - \bar{r}_q)}{\left\{\sum_{i=1}^{n}(r_{pi} - \bar{r}_p)^2 (r_{qi} - \bar{r}_q)^2\right\}^{1/2}}, \qquad \bar{r}_p = \frac{1}{n}\sum_i r_{pi} = \frac{(n+1)}{2} \qquad \bar{r}_q = \frac{1}{n}\sum_i r_{qi} = \frac{(n+1)}{2} \tag{4.6}$$

where r_{pi} and r_{qi} stand for rank value of sequence p and q respectively at location (point) i.

Spearman's rank correlation coefficient also tends to be less sensitive to outliers, compared to the Pearson correlation coefficient.

4.1.5 Mutual information

MI, as a metric based on the random variable entropy measure, was also employed to quantify the relationship between two sequences. MI has some interesting properties where the ability to tackle not only linear but also non-linear properties (unlike correlation coefficient or Euclidean distance) is especially useful [138]. A distance measure derived from MI is given with the

following:

$$\begin{aligned} d_{MI}(p,q) &= 1 - \frac{I(p;q)}{H(p,q)} \\ I(p;q) &= H(p) + H(q) - H(p,q) \\ &= H(p,q) - H(p|q) - H(q|p) \end{aligned} \tag{4.7}$$

where $I(p;q)$ stands for information gain and $H(p,q)$ stands for joint entropy. Further, $H(p)$ and $H(q)$ can be recognized as marginal entropies of variable (sequence) p and q respectively, whereas $H(p|q)$ and $H(q|p)$ are respective conditional entropies. As it can be observed, MI is closely related to the information gain that is commonly used as one of the criteria in decision tree classification models [139].

4.2 Co-activation triggers

Co-activation triggers, as a measure or score, was introduced in this PhD thesis to detect underlying, more subtle changes and relationships that occur among muscle locations in time-localized scope and were counted throughout exercise duration. In this way, co-activation patterns in patients with LBP are quantified and explored. For each channel, two scores were calculated:

1.Co-activation time-alignments of myoelectric activity across all sEMG channels

2.Co-activation misalignments to measure a lack of such time-alignments across all sEMG channels

The procedure was set as follows:

1.Find the prominent localized peaks, for the primary feature being checked, for each channel ch_i, with $i = 1$ to 4.

2.Select one channel and iterate across all maxima (peaks) for that channel. Then, for each iterated peak, with time location loc_t, check in the neighborhood $\pm$ 250 ms if a such prominent peak is also found in other channels. For each channel where such peak is found, increase the channel counter by one in the triggers activation matrix A, $A(ch_i, loc_t) = A(ch_i, loc_t) + 1$.

3.Repeat the procedure for all channels.

4.Co-activation alignment score, for each channel i, is defined as the number of time locations (loc_t) within the triggers activation matrix for which $A(ch_i, loc_t) \geq 3$. Co-activation misalignment score is similarly defined, but for a case where $A(ch_i, loc_t) = 1$, meaning, that the respective time instance of trigger occurred only within that single channel.

5.Normalize resulting co-activation alignment and misalignment scores against exhibited

exercise duration for the respective subject

Altogether, eight secondary features per single primary feature were constructed, $N_{Ca} = 8$.

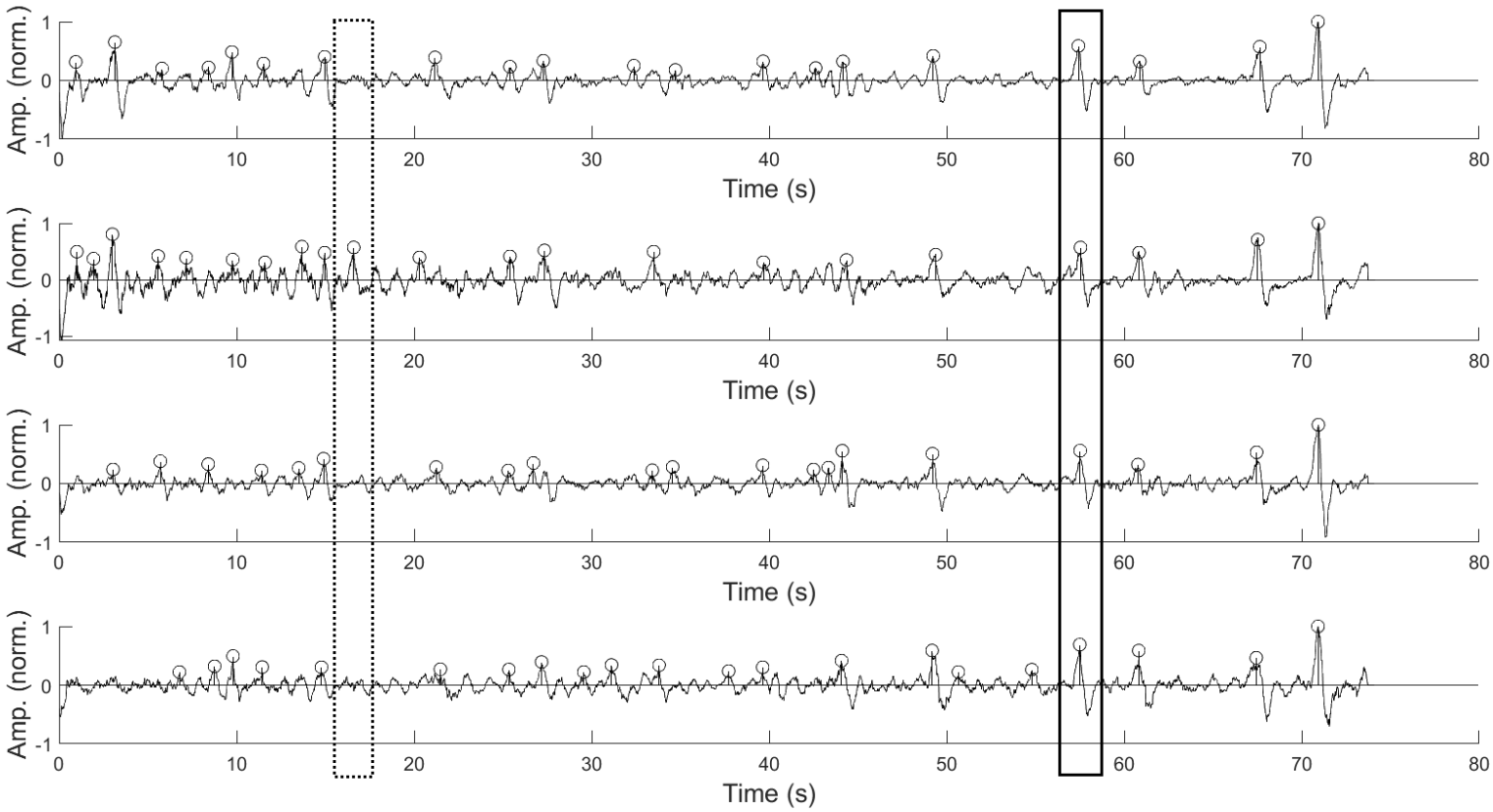


Figure 4.3: Co-activation triggers shown for all four channels, left ULES, left LLES, right ULES, and right LLES, from top to bottom respectively, taken from one healthy subject. The left vertical box, outlined with a dotted line, captures one of the occurrences with missing co-activation alignment (i.e., trigger present only on left LLES muscle location) in $\pm$ 250 ms neighborhood. The right vertical box, outlined with a solid line, captures one of the occurrences with co-activation alignment (i.e., triggers found across all muscle locations). Co-activation triggers were captured via RVD (primary feature) prominent peaks, normalized to max peak value for each channel separately.

4.3 Trends

This feature group tends to establish relationships between the start segment (first ten seconds) and the end segment (last ten seconds) of the static contraction sEMG recordings. Trends are calculated across all primary features separately. Relations are set both within muscles themselves (start vs. end) as well as among muscles location pairs (bilateral and ipsilateral). The rationale behind this approach is at least two-fold: to account for evident intra-channel myoelectric changes as exercise time progresses, and, to account for inter-channel relationships as part of neuromuscular dependencies among muscles (i.e., muscle locations as in our case). Trend features, as part of the secondary features group, are defined as follows:

1. Relative start-end difference for left and right LLES, separately (4.8). Two features are calculated as max and min out of these two metrics for LLES ($N_{Tr_1} = 2$).
2. Relative start-end difference for left and right ULES, separately (4.9). Two features are calculated as max and min out of these two metrics for ULES ($N_{Tr_2} = 2$).
3. Ipsilateral ratio as a relative difference between LLES and ULES end and start segments, for left and right side separately (4.10). Two features calculated as max and min out of these two metrics for left and right ($N_{Tr_3} = 2$)
4. Absolute relative left-right difference calculated for LLES start and end, separately, and normalized with a sum of left and right (4.11). Thus, two relative features per one subject, ($N_{Tr_4} = 2$).
5. Absolute relative left-right difference calculated for ULES start and end, separately, and normalized with a sum of left and right (4.12). Thus, two relative features per one subject, ($N_{Tr_5} = 2$).

The expressions are given with:

$$Tr_1(a) = \frac{LLES_{start}(a) - LLES_{end}(a)}{LLES_{start}(a)} \tag{4.8}$$

$$Tr_2(a) = \frac{ULES_{start}(a) - ULES_{end}(a)}{ULES_{start}(a)} \tag{4.9}$$

where a takes either left or right of the lumbar region. $LLES_{start}$ or $ULES_{start}$ stand for an averaged value of the corresponding (primary) feature for the first 10 s of contraction, whereas $LLES_{end}$ or $ULES_{end}$ stand for the averaged value of the corresponding (primary) feature for the last 10 s of contraction.

$$Tr_3(a) = \frac{LLES_{end}(a) - ULES_{end}(a)}{LLES_{start}(a) - ULES_{start}(a)} \tag{4.10}$$

where a takes either left or right side of $LLES$ and $ULES$ region, separately, for *start* or *end*

segments of the contraction.

$$Tr_4(seg) = abs\left(\frac{LLES_L(seg) - LLES_R(seg)}{LLES_L(seg) + LLES_R(seg)}\right) \tag{4.11}$$

$$Tr_5(seg) = abs\left(\frac{ULES_L(seg) - ULES_R(seg)}{ULES_L(seg) + ULES_R(seg)}\right) \tag{4.12}$$

where *seg* stands for either *start* or *end* segment of the respective sEMG recording, separately, whereas L and R stand for left or right side of these corresponding ULES or LLES regions, respectively.

Altogether, ten secondary features per primary feature were created, $N_{Tr} = 10$, and normalized with respect to the exhibited exercise duration.

A visual inspection of time-evolving changes for the underlying primary features, based on which trends are calculated, is shown in Fig. 4.4. For most of the trends, a time-dependent decrease is observed. For SKEW, PE and RVD, decreasing trend behavior is not explicit or observed in this example.

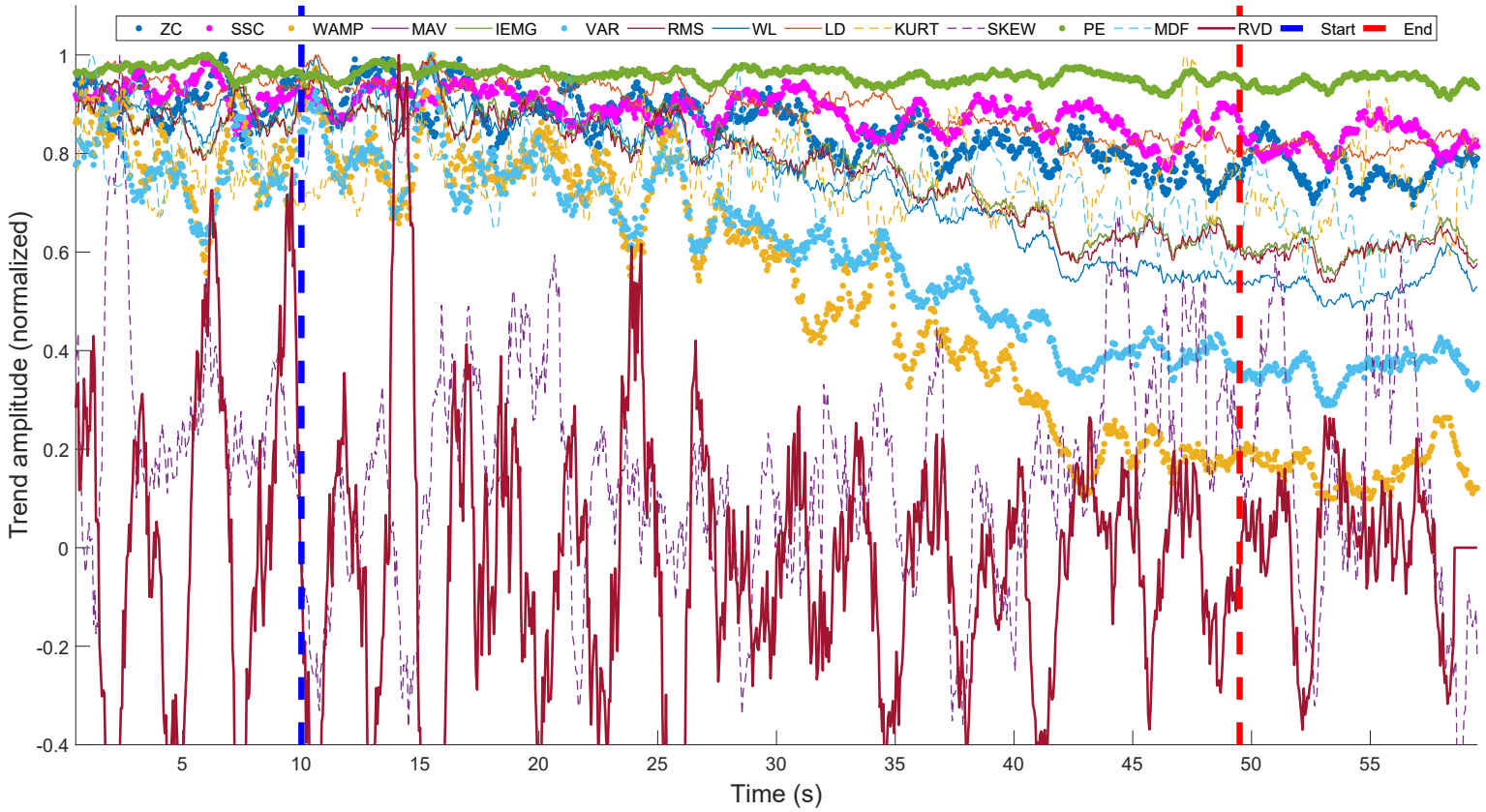


Figure 4.4: Time changes for each of fourteen primary features, calculated from the right LLES muscle location of one CLBP subject. Each primary feature trend is normalized to its max value. Vertical dashed lines represent limits for the start segment (blue), for the first 10 s, and last (red) 10 s of sEMG exercise duration, respectively. Trends for PE (green), SSC (magenta), ZC (blue), VAR (light blue), WAMP (ocher), and RVD (dark red) from top to bottom, respectively, are shown in a thicker line representation.

4.4 Fatigue indices

Muscle fatigue analysis through sEMG has a wide and well-accepted usage [83]. Moreover, muscle fatigue assessment, by exploiting MDF, has been also widely used as an indication for different LBP conditions in patients [31, 54, 92]. In that context, in this PhD thesis a set of extended fatigue indices was included, going beyond previously presented trend features (section 4.3) that also exploit MDF as one of the underlying primary features. Thus, first, the linear regression against MDF time changes was calculated in terms of least mean square error (LMSE) (Fig. 4.5). Then, indices from linear regression slope k_{sl} and initial frequency f_0 were derived (4.13):

$$y_r(x_{MDF}) = k_{sl} \times x_{MDF} + f_0. \tag{4.13}$$

The first subset of MDF fatigue-based features were straightforward linear regression values k_{sl} and f_0, thus eight common features were directly extracted, $N_{FI_0} = 8$ (two features per each sEMG channel).

Considering different absolute values for MDF, not only among different subjects but also among sEMG channels for the same subject (especially valid for the derived parameter f_0), additional fatigue indices were composed to reflect the relative nature between k_{sl} and f_0 parameters as fatigue is progressing. These relative indices were defined as follows:

1. Relative left-right side difference for k_{sl} for ULES and LLES, respectively, normalized with the higher of k_{sl} value between left and right. Thus, two relative features ($N_{FI_1} = 2$).
2. Relative left-right side difference for f_0 for ULES and LLES, respectively, normalized with the higher of f_0 value between left and right. Thus, two relative features ($N_{FI_2} = 2$).
3. Relative left-right side difference for k_{sl}/f_0 for ULES and LLES, respectively, normalized with the higher of k_{sl}/f_0 value between left and right . Thus, two relative features ($N_{FI_3} = 2$).
4. Left-right side ratio for $f_{0,L}/f_{0,R}$ for ULES and LLES, respectively. Same relative left-right ratio for $k_{sl,L}/k_{sl,R}$ was calculated. Thus, four relative features ($N_{FI_4} = 4$).
5. Up-down ipsilateral ratio for $f_{0,U}/f_{0,D}$ for left side and right side, respectively. Same relative up-down ratio for $k_{sl,U}/k_{sl,D}$ was calculated. Thus, four relative features ($N_{FI_5} = 4$).
6. Ratio k_{sl}/f_0 for all four channels separately. Thus, four relative features per subject ($N_{FI_6} = 4$).

Altogether, this feature group provided a total of twenty-six features, $N_{FI} = 26$.

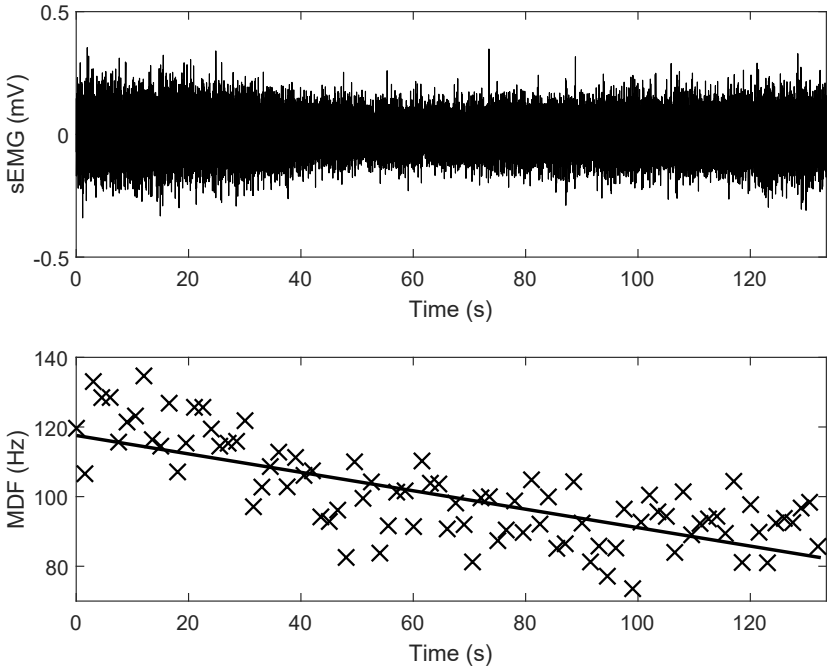


Figure 4.5: Raw sEMG signal (top) acquired with FREEEMG measurement system (section 3.1.2). Linear regression (solid line) calculated LMSE-wise with respect to MDF values (crosses). Regression line slope corresponds to k_{sl} and intersection with y-axis corresponds to the initial frequency f_0.

4.5 Overall features set

In this step, all calculated features are gathered to create a feature vector for each subject at hand. It is important to notice that each of the fourteen primary features (Table 3.2) are contributing to corresponding secondary feature groups presented in this chapter, except for fatigue-related indices where only MDF feature is used. As a result, the number of variables per single observation is multifold increased. Accordingly, the number of features per secondary feature group is as follows:

1. Coordination measures: ten secondary features per single primary feature ($N_{Co} = 10$), altogether $N_{sg1} = N_{Co} \times 14 = 140$
2. Co-activation triggers: eight secondary feature per single primary feature ($N_{Ca} = 8$), altogether $N_{sg2} = N_{Ca} \times 14 = 112$
3. Trends: ten secondary features per single primary feature ($N_{Tr} = 10$), altogether $N_{sg3} = N_{Tr} \times 14 = 140$
4. Fatigue-related indices: twenty-six secondary features calculated directly from MDF, as presented in section 4.4, altogether $N_{sg4} = N_{FI} = 26$

On top of this, for coordination and co-activation features specifically, the autocorrelation-based features were also created. More concretely, before applying distance measures (coordination) or triggers detections (co-activation), (time)autocorrelation for each primary feature

was calculated following the expression:

$$r_k = \frac{c_k}{c_0},\ c_k = \frac{1}{T}\sum_{t=1}^{T-k}(y_t - \bar{y})(y_{t+k} - \bar{y}) \tag{4.14}$$

where k represents a lag (in our case a full shift across the whole sequence length was used), and c_0 stands for the variance of the given time sequence y_t. This way, the number of secondary features for coordination and co-activation groups is doubled.

Additionally, for comparison and validation reasons, the overall secondary features set was expanded with single-valued representations of primary features, in a simplistic manner. Namely, each primary feature time sequence (calculated directly from the raw sEMG signal as presented in section 3.2) was contributing with only one value, and that is, the maximum value calculated on the whole sequence (mean, median, 95th, and 99th percentile values were also explored). Thus, a macroscopic representation of each primary feature was included, but at the cost of losing insight into time-related dynamics of myoelectric activity. This approach resulted with four additional features (one per channel) per each primary feature, altogether $N_P = 4 \times 14 = 56$ features.

As a last contribution to the overall feature set, the time length (T_{emg}) of the usable portion of sEMG recording exploited in this research, extracted for each subject, is added as additional feature [9, 31], thus one feature per subject ($N_T = 1$).

Finally, the total number of features is the following sum $N_{sg,Tot} = 2 \times (N_{sg1} + N_{sg2}) + N_{sg3} + N_{sg4} + N_P + N_T = 727$.

Chapter 5

Classification modeling and features selection

In this chapter, the procedure of analyzing and selecting the optimal subset of primary features, by means of correlation and statistical analysis, is presented. This led to the next step of selecting features for the secondary features groups, where *neighborhood component analysis* (NCA), as the main technique for feature selection, was introduced. Such an approach enables a significant reduction in the number of features to tackle, together with preserving the contextual information in the original domain.

To validate the feature selections and models definition, different classification models are employed in order to verify the proposed method based on sEMG with LBP-specific features that were previously constructed. The intention was to establish a meaningful set of features that can be correlated with actual subjects' health conditions expressed via myoelectric activity, triggered by endurance exercise and related underlying neuromuscular processes, with or without the presence of LBP.

Subsequently, as the first validation step, the best selected individual primary features were validated against the available classification models. Next, the resulting primary features were grouped and used to calculate the respective contextual (secondary) feature sets. These secondary feature sets were then additionally compared to several other calculated contextual feature sets, that were based on primary features proposed by other authors and studies (namely, Hudgin's and Du's vectors for comparison). This provided insights and helped in understanding whether our selected primary features (SSC, WAMP, VAR, ZC, RVD + PE) were a satisfactory choice. Further, to enable an insight into which contextual feature group (Coordination, Co-activation, Trends, and Fatigue frequency-related indices) individually contributes to the classification accuracy results most, and to which extent, these groups were compared among each other.

Finally, two-class classifications and differentiations among three subject groups at hand

(HS, CLBP, RLBP) were performed, for the following classification pairs: (I) HS vs. LBP, (II) HS vs. RLBP, (III) HS vs. CLBP, (IV) CLBP vs. RLBP as presented in Table 5.5.

The respective data set at hand consisted of ninety-one subjects, divided into three groups (HS, CLBP, RLBP), with a share of 40.7%, 31.9%, and 27.4% among groups, respectively. Demographic and biometric statistics did not point to any significant deviation or differences among groups when considering age, height, weight, and BMI (Table 3.1).

5.1 Feature analysis and selection

5.1.1 Primary features relationships analysis

The initial set of fourteen primary features led to many secondary features ($N_{Tot} = 727$). In the next step, the linear relationships were checked by calculating the correlation for each pair of primary features.

$$\begin{bmatrix} f_{1,1,1} & f_{2,1,1} & & f_{N_f,1,1} \\ f_{1,1,2} & f_{2,1,2} & \cdots & f_{N_f,1,2} \\ \vdots & \vdots & & \vdots \\ f_{1,1,n_1} & f_{2,1,n_1} & & f_{N_f,1,n_1} \\ f_{1,2,1} & f_{2,2,1} & & f_{N_f,2,1} \\ f_{1,2,2} & f_{2,2,2} & \cdots & f_{N_f,2,2} \\ \vdots & \vdots & & \vdots \\ f_{1,2,n_2} & f_{2,2,n_2} & & f_{N_f,2,n_2} \\ \vdots & \vdots & \vdots & \vdots \\ f_{1,N_s,1} & f_{2,N_s,1} & & f_{N_f,N_s,1} \\ f_{1,N_s,2} & f_{2,N_s,2} & \cdots & f_{N_f,N_s,2} \\ \vdots & \vdots & & \vdots \\ f_{1,N_s,n_{N_s}} & f_{2,N_s,n_{N_s}} & & f_{N_f,N_s,n_{N_s}} \end{bmatrix} \tag{5.1}$$

For this correlation analysis, the intention was to preserve the aspect of time progression for the intrinsic characteristics of respective feature sequences. Therefore, for each primary feature sequence, a new vector was composed by stacking individual feature sequences for each

subject, one on top of the other. Such vertically stacked vectors, for each feature, were then stacked column-wise into the matrix for all primary features at hand. The matrix composition is given with expression (5.1), where indices triplet (i, j, k) for each matrix element corresponds to i-th primary feature (N_f represents the total number or features), j-th subject s (N_s represents a set of subjects, either related to healthy subjects or subjects with LBP, only) and k-th element in the time sequence of the respective primary feature with sequence length n_{N_s} for subject s (all feature sequences for same subject are of the same length). The feature vector for each subject is standardized ($\mu = 0$, $\sigma = 1$) before being added to the matrix to eliminate the effects of different absolute value ranges among the same features across different subjects.

The correlation coefficient was calculated using Pearson correlation as given by

$$\rho(a,b) = \frac{\sum_{i=1}^{n} (X_{a,i} - \bar{X}_a)\left(Y_{b,i} - \bar{Y}_b\right)}{\sqrt{\sum_{i=1}^{n} (X_{a,i} - \bar{X}_a)^2 \left(Y_{b,i} - \bar{Y}_b\right)^2}} \tag{5.2}$$

where X_a and Y_b stand for column-wise stacked feature vectors composed of either set or healthy or set of LBP-related subjects, only.

Subsequently, correlation coefficients were calculated for HS or LBP subjects, separately, thus avoiding potential mixing issues or masking the characteristics pertained to each of the (sub)groups individually. This resulted in two groups of correlation matrices (14×14, $N_f = 14$), one for HS and one for LBP subjects, as given in Fig. 5.1 and Fig. 5.2. Each group has four correlation matrices, one for each channel, this way providing valuable insight with Pearson correlation ρ and p-values using Student's t-distribution for a transformation of the correlation [140].

Some level of correlation among all feature pairs was detected with p-values $p < 0.05$, i.e. not for a single pair of features it can be stated with confidence >95% that $\rho = 0$, although, for most pairs, it was relating to low-level correlation. However, a very strong correlation, $\rho^2 > 0.8$, has been observed for the following feature subsets across all channels, and irrespective of the subject's health condition state (HS vs. CLBP vs. RLBP):

1.MAV, IEMG, LD, RMS, VAR

2.SSC, PE

In order to provide a more complete statistical analysis applied to the primary features set, additionally, ANOVA was performed by examining the impact of two effects (two-way ANOVA). Having in mind that responses in this analysis are not univariate metric values, but earlier mentioned time sequences (per each given primary feature), the analysis had to be adjusted and simplified in a way to represent the feature time sequence with only one value. This feature value was selected to be the max value in the respective sequence. The effects of two factors on the response of the dependent variable were analyzed. These two effects were: the ef-

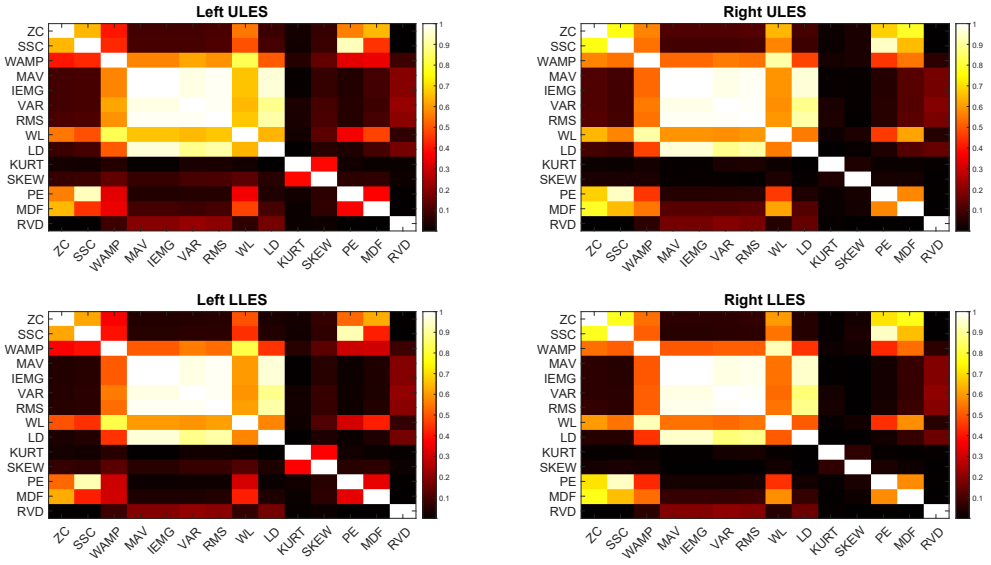


Figure 5.1: Correlation matrix for all fourteen primary features (ZC, SSC, WAMP, MAV, IEMG, VAR, RMS, WL, LD, KURT, SKEW, PE, MDF, RVD) for a set of healthy subjects (HS) only, and each sEMG channel, left and right ULES (top), and left and right LLES (bottom). Correlation values are shown as squared Pearson correlation coefficient, ρ^2.

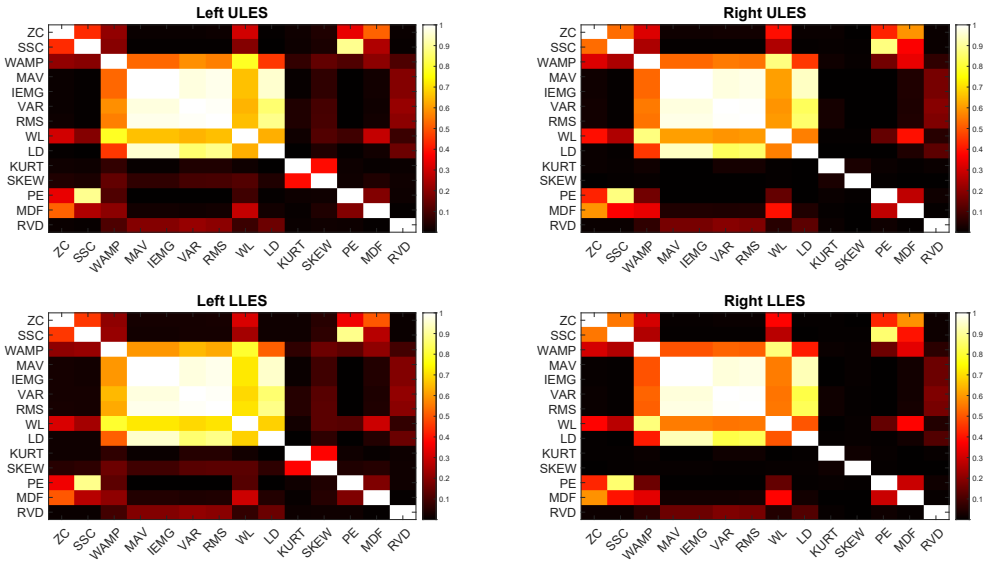


Figure 5.2: Correlation matrix for all fourteen primary features (ZC, SSC, WAMP, MAV, IEMG, VAR, RMS, WL, LD, KURT, SKEW, PE, MDF, RVD) for a set of LBP subjects (CLBP and RLBP) only, and each sEMG channel, left and right ULES (top), and left and right LLES (bottom). Correlation values are shown as squared Pearson correlation coefficient, ρ^2.

fect of selecting one of the primary features (Table 3.2), and the effect of selecting the subjects' health status categorized as HS or LBP. The statistical significance ($p < 0.05$) is shown for the interaction of these two effects for left and right LLES muscle sites, whereas for ULES muscle sites such interaction with clear statistical significance was not obvious. Additionally, for some of the muscle sites, the mean values of the analyzed representative max-values for WAMP, WL, and RVD are shown to be different between LBP and HS subjects with statistical significance ($p < 0.05$).

5.1.2 Feature selection for secondary groups

Neighborhood component analysis (NCA) was selected as the main technique for feature selection in this PhD thesis [141]. NCA also demonstrates the capability of reducing irrelevant or redundant features while keeping the selected feature vectors in the original vector space, unlike some other commonly used feature reduction or feature projection techniques (e.g. principal component analysis - PCA or linear discriminant analysis - LDA) (Fig. 5.4). This approach opened up a possibility for getting a better insight into understanding which features and properties have the key impact on the LBP detection and differentiation among subjects.

NCA method, implicitly, throughout the procedure of maximizing the objective function, measures the average leave-one-out classification accuracy and sets respective weights (w) for individual features [142]. The classification method used behind is a variant of the 1-NN classifier. The aim is to find such weights vector w that will optimize the nearest neighbor classification accuracy. In the context of weights vector w, a weighted distance between two samples, of dimension d, is defined with

$$D_w\left(x_i, x_j\right) = \sum_{l=1}^{d} w_l^2 \left|x_{il} - x_{jl}\right|. \tag{5.3}$$

To maximize the leave-one-out classification accuracy by employing the nearest neighbor, a classification accuracy function in a differential form is required. Knowing that the leave-one-out form is not a differential function *per se*, we are looking into the approximation of the probability distribution function with the following expression

$$p_{ij} = \begin{cases} \frac{\kappa\left(D_w(x_i,x_j)\right)}{\sum_{k\neq i}\kappa(D_w(x_i,x_k))}, \text{ if } i \neq j \\ 0, \text{ if } i = j \end{cases} \tag{5.4}$$

where $\kappa(z) = exp(-z/\sigma)$ is a kernel function of kernel width σ. Kernel function was chosen in such a way to secure that for small distances $D_w(x_i, x_j)$, the resulting probability p_{ij} (to have sample x_i choose x_j as a reference point) is set high. Two characteristic edge cases for kernel width σ can be observed:

- $\sigma \to 0$ where only the closest neighbor can be selected as a reference point
- $\sigma \to +\infty$ where all points have equal chance to be selected (except x_i point)

Following (5.3) and (5.4), the probability to have point x_i correctly classified is

$$p_i = \sum_j y_{ij} p_{ij} \tag{5.5}$$

where $y_{ij} = 1$ if and only if $y_i = y_j$, otherwise $y_{ij} = 0$. Therefore, the approximate classification accuracy can be written as

$$\xi(w) = \frac{1}{N} \sum_i p_i = \frac{1}{N} \sum_i \sum_j y_{ij} p_{ij} \tag{5.6}$$

with N number of observations, where for $\sigma \to 0$, $\xi(w)$ becomes exact leave-one-out classification accuracy. Furthermore, to enable the feature selection (and prevent classification model overfitting), a new regularization factor is introduced resulting in the subsequent objective function:

$$\xi(w) = \sum_i \sum_j y_{ij} p_{ij} - \lambda \sum_{l=1} w_l^2 \tag{5.7}$$

where λ stands for the regularization parameter which can be further fine-tuned throughout an iterative process with cross-validation. Furthermore, it is also good to note that only one regularization parameter λ is used for all the weighted factors w_l, thus bringing less complexity in finding an adequate value.

Now, when having a differentiable (approximate) function $\xi(w)$, we are searching for such w vector for which $\frac{\delta \xi(w)}{\delta w_l}$ will provide maximum value. This search is done by employing some of the iterative numeric gradient optimization methods, like the proposed gradient ascent technique [142, 143]. It is shown that the given optimization procedure, by including the regularization parameter, results in many weighted factors $w_l \to 0$. Features with such weights are excluded, thus effectively reducing the number of features significantly. In this research, a threshold of 0.02 was applied as an exclusion criterion. The remaining features are forming the selected features subset for subsequent classification steps.

An example of NCA feature selection based on resulting feature weights is given with Fig. 5.3, where the resulting twenty features were selected out of the initial 327.

5.2 Classification models

Several different groups of classifier types and classification models were examined in combination with the constructed feature sets. *Classification Learner* application as part of Matlab R2020b (The Mathworks, Inc., Natick, Massachusetts) was used as a software package [144]

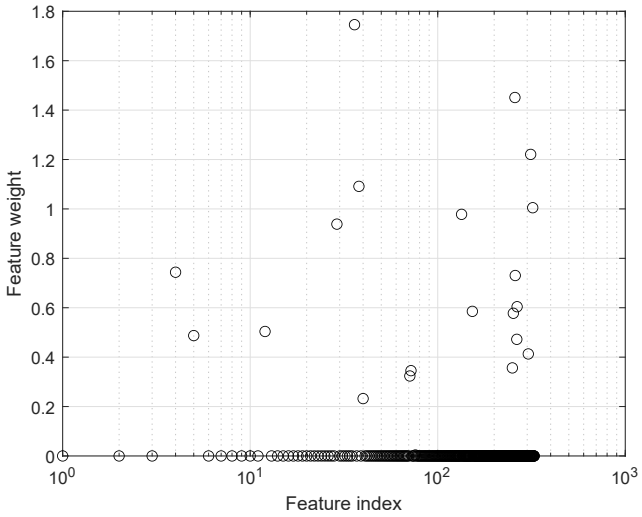


Figure 5.3: Example of feature weights for each individual secondary feature as part of the NCA feature selection procedure for a submitted secondary feature set based on SSC, WAMP, VAR, ZC, RVD, and PE primary features. The secondary feature set consisted of 327 features. Only features with feature weights > 0.02 were included in the resulting subset of features. Twenty features were selected this way.

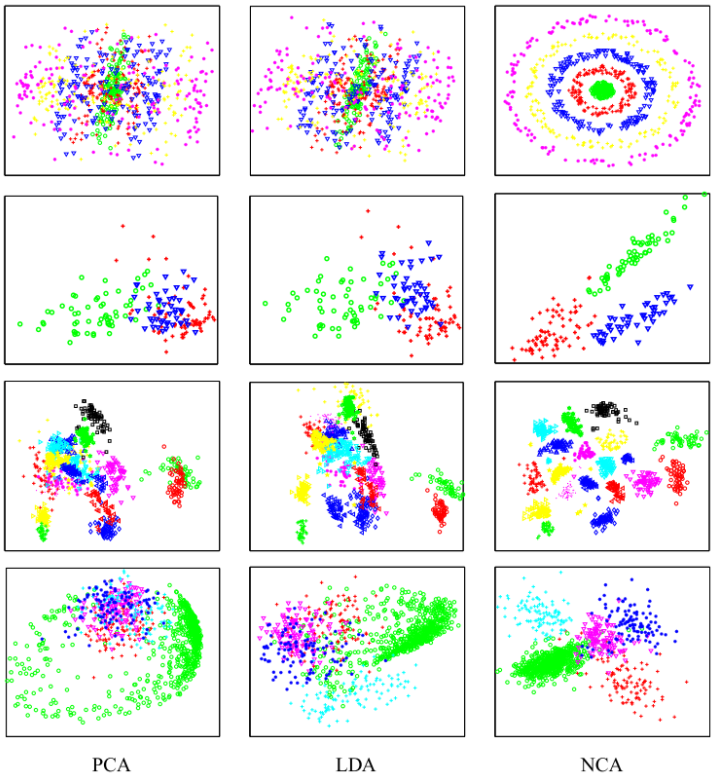


Figure 5.4: Comparison among different methods (PCA, LDA, NCA) for high-dimensional feature vectors reduction to two-dimensional vector space, applied on a different domains' problems: "concentric rings", "wine", "faces" and "digits" datasets, from top to bottom. Picture with example is taken from [141].

to prepare data, select classification models, conduct classification procedures and validate respective models. The predefined classification models were chosen *as-is* without any additional parameters optimization or fine-tuning steps. Thus, the goal was to primarily validate the potential of the proposed approach with LBP domain-related features constructed as previously discussed. Matlab classification models in this research, used *out-of-the-box*, can be grouped into the following:

- •Decision trees (DT) with fine, medium, and coarse splits. Strong foundations for this approach and its popularity were established by L. **B**reiman, J. **F**riedman, R. **O**lshen i C. **S**tone (BFOS group) [145].
- •Discriminant analysis: linear (LDA) and quadratic (QDA). Based on Fisher's discriminant analysis [146, 147].
- •Logistic regression (LR) [148]
- •SVM with different kernel types (linear, quadratic, cubic, Gaussian) and with different scales for Gaussian kernels (fine, medium, and coarse). [149, 150].
- *k*-nearest neighbor (*k*NN) with fine, medium and coarse resolution depending on parameter (*k*) and different distance metrics (Euclidean, cosine or cubic), with or without a weighted distance [151, 152].
- •Ensemble classifiers based on *bagging* (developed by Breiman *et al.* [153]), *boosting* algorithms as developed by Schapire, Friedman *et al.* [154, 155, 156] and *random subspace* techniques [157, 158].

A given list resulted with twenty-three different classification model variants used to validate selected feature sets and proposed method overall as listed in Tables 5.1, 5.4, 5.5.

5.2.1 Decision trees

DT algorithms are based on the recursively repeated splits of data set by applying the selected criterion that maximizes the separation of the data, which in turn creates a tree-like structure [145]. Most commonly used separation criteria are information gain and Gini's index [139, 145], somewhat less commonly used cross-entropy (also known as maximum deviance reduction) [159], however, all being entropy-based measures. The splitting objective is to minimize the dissimilarity in the terminal nodes by maximizing a decrease in entropy due to the split (i.e., selecting the split with the highest information gain or highest decrease in node impurity). The classification process results in an estimate of $P(y|x)$ ratio of y class elements over all elements of the leaf node that contains data sample x to be classified.

A classic CART algorithm [145] was employed with the predefined splitting criterion (Gini's diversity index) and splitting resolutions (namely, fine, medium, and coarse with 100, 20, 1, as a maximum number of splits, respectively).

Compared with the other machine learning methods mentioned here, decision trees have the

advantage that they are not black-box models, but can easily be expressed as a set of rules. In many application domains, this advantage weights more heavily than the drawbacks, so these models are widely used in medicine. [148]

5.2.2 Discriminant analysis

LDA and QDA are classification and dimensionality reduction techniques that are developed based on Fisher's linear discriminant. These methods tend to express a dependent variable (sample class) as a linear combination of other independent features. The classification objective is to find a such linear combination that best separates two or more classes [147]. Thus, the class label to be explained is a category inferred by a linear combination of independent continuous feature vectors.

LDA classifier requires certain assumptions to perform well. These assumptions are normally distributed classes and equal class covariances. In such cases, quite an intuitive concept of discriminant analysis via Fisher's linear discriminant rule can be given by

$$S = \frac{\sigma^2_{between}}{\sigma^2_{within}} \tag{5.8}$$

where S represents a separations criterion between groups (classes). Meaning, a separation between classes is achieved by finding such a combination of linear predictors that result in maximum variance between the classes ($\sigma^2_{between}$) and minimum variance within the same class samples (σ^2_{within}) [146].

In cases when the assumption for equal class covariances is violated, QDA is proposed to be employed. QDA can deal with more complex separation problems (quadratic separation surface) with no equal covariances assumption as a constraint. However, this comes at the cost of more computational complexity and significantly more parameters to be estimated - as now each class requires its own covariance matrix [160].

Classification models employed in this research were utilizing a full covariance matrix structure as a predefined value for both LDA and QDA classifiers (as compared to another diagonal-only option).

5.2.3 Logistic regression

Logistic regression is a type of statistical model that employs a logistic function (sigmoid function) to construct probability models for binary dependent variables. Such probability models can be turned into classifiers by introducing a threshold value where input values (e.g. feature vectors) with a probability above the threshold are classified as one class, and if below, as the other class. In that way, logistic regression calculates the class membership probability for one

of the two categories in the data set [148]. This probability can be given in a simplified form as

$$P(1|x,\alpha) = \frac{1}{1+e^{-(\alpha \cdot x)}} \tag{5.9}$$

and $P(0|x,\alpha) = 1 - P(1|x,\alpha)$, where α stands for model parameter values that need to be estimated, by maximizing the maximum likelihood estimation $\Pi_{i=1}^{n} P(y_i|x_i,\alpha)$ in search for optimal values α [148, 161].

No other parameters were required to be set or defined for this classification approach and it can be extended to multiclass problems as well [162].

5.2.4 Support vector machines

SVM is one of the most popular and most used machine learning and classification techniques. It has been shown to be very robust and among the most accurate methods based on good mathematical and theoretical foundations. Some of the key strengths are the ability to learn on a limited number of training samples and to successfully deal with high feature attributes dimensionality (thus, performance independent of a number of dimensions) [139].

Initially, SVM was constructed for two-class classification problems, but it is possible to extend it to multiclass classification problems as well [163]. In the two-class classification examples, the procedure's idea is to find the best classification function that enables differentiation between two data sets with the realization (interpretation) in a geometric sense. Namely, for a data set that can be separated linearly, a linear classification function corresponds to a separating hyperplane $f(x)$ that runs through the middle between these classes, thus separating them. After such function is defined, any new sample x_n can be classified in a simple way by examining the resulting value sign for respective definition $f(x_n)$, where x_n is assigned to "positive" class if $f(x_n) > 0$.

However, taking into consideration that there are many hyperplanes with the required properties, the condition that ensures finding the best separating function is set by maximizing the area of separation, i.e., margin, between two groups. Geometrically, the margin corresponds to the shortest distance between the closest sample in the data set and the hyperplane (Fig. 5.5). Such an approach enables the generalization of the problem by ensuring not only the best classification results on the learning data set but also leaving the room (i.e., margin) for the new (test) data. The hyperplane margin maximization task is achieved by minimizing the following function form with respect to w and b:

$$\min L_P = \frac{1}{2}\|w\| - \sum_{i=1}^{t} \alpha_i y_i (w \cdot x_i + b) + \sum_{i=1}^{t} \alpha_i \tag{5.10}$$

where x stands for a learning sample (observation), y indicates the class (-1 or 1 in this case)

that x belongs to, t stands for the number of learning samples, $\alpha_i = 1,...,t$ are such non-negative numbers so that the derivation of L_P in respect to α_i equals zero. α_i are in fact Lagrange multipliers, and L_P is called *Lagrangian*. Vectors represented by w and a constant b are the parameters that define the hyperplane.

Next, this analytical problem for the so called *primal* form L_P is translated into *dual* form notation L_D

$$\max L_D = \sum_{i=1}^{t} \alpha_i - \frac{1}{2} \sum_{i=1}^{t} \alpha_i \alpha_j y_i y_j \left(x_i \cdot x_j\right) \tag{5.11}$$

where the dependency on w and b has been removed and the respective solution is found as the maximum in respect to Lagrange multipliers α_i, with the conditions set as:

$$w = \sum_{i=1}^{t} \alpha_i y_i x_i, \ \sum_{i=1}^{t} \alpha_i y_i = 0 \ \& \ \alpha_i \geq 0 \tag{5.12}$$

This formulation enables finding a such solution by only calculating the inner product $x_i \cdot x_j$ which has been shown to be very useful in expanding the problem from only linearly separable data set to general (non-linear) cases. In such a case, the optimization problem becomes

$$\max L_D = \sum_{i=1}^{t} \alpha_i - \frac{1}{2} \sum_{i=1}^{t} \alpha_i \alpha_j y_i y_j K\left(x_i, x_j\right), \ K\left(x_i, x_j\right) = \Phi(x_i) \cdot \Phi(x_j) \tag{5.13}$$

by transforming the data from one multidimensional space into another $\Phi : x \rightarrow \phi(x)$, in which we believe the data would be easier to separate by the resulting hyperplane (Fig. 5.6)

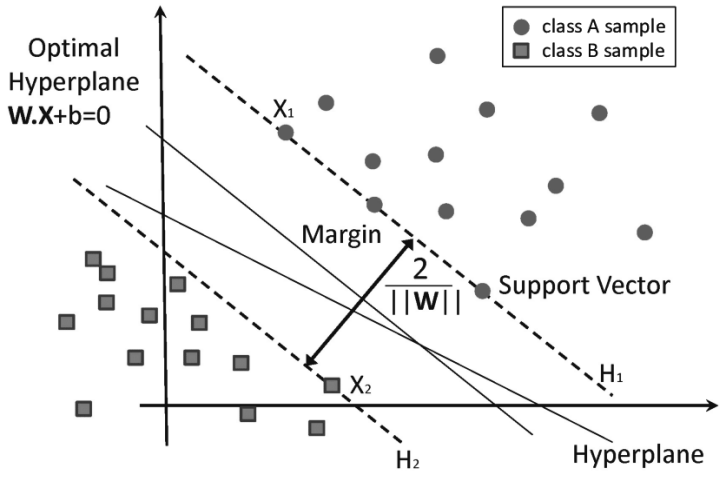


Figure 5.5: Optimal hyperplane for the maximuim margin and respective margin areas for the case with two separable classes. Samples placed at the edges (margins) are called the support vectors. Picture is taken from [164] and adjusted under license given by `https://creativecommons.org/licenses/by/4.0/`.

For the given transformations, different *kernel functions K* are utilized to replace the inner product $x_i \cdot x_j$. This way, by choosing the appropriate kernel functions, different non-linear

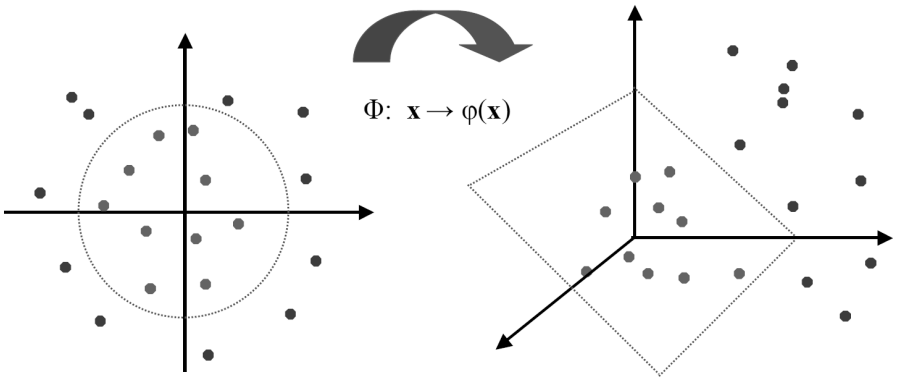


Figure 5.6: Data transformation from one multidimensional vector space into another by applying non-linear kernel functions. Picture is taken from `https://www.cs.umd.edu/~samir/498/SVM.pdf` and adjusted.

transformations can be achieved which enables better flexibility in finding optimal solutions for the respective data sets at hand. Some of the kernel functions K types can be defined with:

- $K(x_i, x_j) = (x_i \cdot x_j + 1)^p$, where, by selecting value for parameter p ($p = 1$, 2 or 3), a linear, quadratic or cubic kernel is defined, respectively.
- $K(x_i, x_j) = \exp\left\{-\|x-y\|^2 \Big/ 2\sigma^2\right\}$, where the exponential function represents the radial basis, i.e. Gaussian kernel function

In this setup, auto-selected values, as per [144], for kernel widths (Gaussian kernel function) were applied for three different resolution variants: fine, medium, and coarse resolution. Soft-margin, as a predefined separation criterion, was applied to all SVM classifiers variants employed in this thesis. [149, 150].

5.2.5 *k*-nearest neighbor

k-NN is a simple classification algorithm that can perform well in many scenarios, including multiclass classification cases. Key elements that determine the outcome of such a nearest neighbor approach are (A) selection of value k, (B) distance or similarity measure, and (C) the structure of the labeled set. The choice of k, if set too small, can result in models being sensitive to noise points. On the other hand, for large k, a risk of including too many points from other classes in the neighborhood is imminent [139]. As for distance measures (B), a key principle is to engage such metric for which a smaller distance between two objects implies a greater likelihood of having the same class. Some measures, like Euclidean distance, have difficulties dealing with the high dimensionality of data and become less discriminating as the number of attributes increases. Related to this, problems with one attribute dominating distance measure, over other ones, can also be observed. Therefore, different scaling and feature weighting approaches were introduced [165]. Additionally, the approach of combining class labels, in the given neighborhood for the sample at test (i.e. calculating probability $P(y|x)$ that class y was de-

tected at sample location x), is another aspect that requires attention, where, in a certain number of cases, a simple majority voting (counting) does not suffice.

In this research, predefined k values 1, 10, and 100, were tested, resulting in fine, medium, and coarse resolution, respectively (with predefined Euclidean distance used). Also, two other distance metrics were exploited, cosine and cubic, with medium resolution ($k = 10$). Finally, a variant with square-inverse weighted distance measure was employed as well (on Euclidean distance), instead of simple majority voting (counting) approach.

k-NN classifiers are lazy learners, that is, models are not built explicitly unlike eager learners (e.g., decision trees, SVM, etc.). Thus, building the model is cheap, but classifying unknown objects is relatively expensive since it requires the computation of the k-nearest neighbors of the object to be labeled. [139]. This type of classifier also poses a challenge with defining appropriate metrics to measure between data items.

5.2.6 Ensemble learning

Ensemble learning is a term for a general technique, or a concept, of employing multiple learners (multiple instances of a classifier) to solve a classification problem. The main idea (or hypothesis) behind driving such a concept was that ensemble generalization ability is significantly better than of a single learner, and thus can perform better. Namely, one of the crucial considerations was whether a weak learning algorithm (performing just slightly better than a random guess) could be *boosted* in such a way to result in equally accurate outcomes as when strong learning algorithms were used [139]. Robert E. Schapire provided theoretical background by proving that such construction is possible, thus effectively bringing in the first Boosting algorithm [154]. In this PhD thesis boosted trees classifier, based on the AdaBoost algorithm was utilized [155]. Additionally, another variant of boosting ensemble classifier was also used, namely, **R**andom **U**nder **S**ampling (RUS) boosting trees [158]. A key advantage is the effectiveness of classifying imbalanced data, in cases some class in the training data set has many fewer members than another.

Leo Breiman has contributed strongly to the construction of *bagging* (bootstrap aggregating) [153] and *random forests* [166] algorithms. In this PhD thesis bagging was used where, in a nutshell, it is a method where multiple versions of a classifier are generated by bootstrap replicates (i.e., sampling with replacement) of the learning set and using these as new learning sets [153]. The overall classifier is constructed by combining classification results in a majority voting manner.

Further, random subspace methods were employed where the ensemble classifier was constructed by a series of (weak) learners that operate on a subset (thus *subspace*) of randomly selected features [157]. As ensemble (weak) learners, different classifiers can be employed. In this research, (linear) discriminant and k-nearest neighbor classifiers were employed. Classifi-

cation prediction outcome was constructed by taking an average of the score prediction of the weak learners, and classifying the category with the highest average score [144].

As in the previous cases, predefined out-of-the-box parameter values were applied, where for all ensemble classifiers the number of learners was set to 30, with a learning rate of 0.1. Additionally, for boosted algorithms (boosted trees and RUSBoosted trees), the maximum number of splits was set to 20, whereas for bagged trees it was set to 90. Gini's diversity index, as a default split criterion, was used in both cases.

5.3 Classification results and models verification

So far, the proposed method for raw feature extraction, novel feature construction, and feature selection were presented. The following section is presenting verification with the aim to: (A) confirm the choice of features selected and (B) apply the feature construction outcomes to the classification models in order to detect LBP patients and consequently make differentiation among subject groups at hand (HS, CLBP, RLBP).

Features verification process (A) was established by performing a series of classifications and cross-validations with comparisons, utilizing two-class classifiers. Therefore, the data set (ninety-one subjects) was initially divided into two groups, namely, healthy subjects (HS) and LBP patients (CLBP and RLBP combined), with thirty-seven and fifty-four subjects, respectively (Table 3.1). Following feature validation outcomes in (A), a series of two-class classifications (B) among three groups at hand was performed, separately. More concretely, two-class classifications were performed for the following differentiations among groups: HS vs. CLBP, HS vs. RLBP, CLBP vs. RLBP. As a basic step, the overall HS vs. LBP differentiation was performed, where general separability between LBP patients and non-LBP subjects was tested.

5.3.1 Cross-validation

Feature selection (based on NCA) and models verification (twenty-three classifiers) procedures were performed on a full data set in an M-fold cross-validation manner. Meaning, no separate untouched test set was used, but the data set at hand was divided into M nonoverlapping subsets of equal size, where the prediction model was trained on $M-1$ folds combined together, whereas the prediction accuracy (i.e. error) was estimated on the remaining samples (single fold). Cross-validation was opted to make full use of all LBP patient samples having in mind the specific LBP domain challenges and the expected heterogeneity among the patients. This approach is relatively common in cases when the number of samples in the data set is relatively small. However, this approach introduces a certain selection bias that generally can result in a too-optimistic estimate of the prediction error rate [167, 168].

For the NCA feature selection procedure, the number of folds was equal to the number of samples in the data set, thus leave-one-out cross-validation was performed. Leave-one-out cross-validation is considered to be nearly unbiased, but with potentially high variance. For the verification of models utilizing twenty-three classifiers, a cross-validation procedure with ten-folds ($M = 10$) was performed for every single classification. Further, each classification with cross-validation was repeated ten times ($N_{iter} = 10$). Finally, all results were averaged. In this way, more statistical confidence by reducing the effects of results' variations was introduced.

Alongside classification accuracy (ACC), the values for sensitivity (recall, RC) and precision (positive predictive value, PPV) were also calculated in the same manner as a result of multiple iterations and averaging.

Additionally, each classification experiment was conducted fully independent, thus different classification comparisons (Table 5.3, 5.4, 5.5) that were employing the same feature sets and models (e.g., feature sets relating to Top 5+PE), could have resulted in slightly different results in different experiments, due to statistical variation, but at the same time giving the notion of results consistency.

5.3.2 Primary features selection verification

To confirm the contribution of each individual primary feature to the overall contextual feature set (as defined in section 4.5), and to relate the results with earlier performed correlation analysis (section 5.1.1), the verification across all classifiers was performed.

The verification across all primary features and all classifiers was conducted for HS vs. LBP two-class classification, as follows:

1. For each primary feature, the corresponding secondary features set was calculated (based only on that single primary feature)
2. NCA procedure for features selection was performed (thus, reducing the number of features we are working with)
3. Each of twenty-three classification models was employed by taking previously NCA selected features set as an input
4. Cross-validation procedure with ten folds ($k_{fold} = 10$) was performed, as presented in section 5.3.1

Results for this intermediate verification step are presented in Table 5.1.

For each primary feature, an averaged classification accuracy rank was calculated across all classifiers employed to provide an insight into how well the respective feature resonated with the given secondary feature groups' modeling. These ranks are given in Table 5.2.

SSC performed best for nine classifiers (especially for decision trees), whereas PE performed best for eleven classifiers (especially for SVM variants), out of twenty-three classifiers (and variants) employed. Among all classifiers and all features, SSC provided the single best

Table 5.1: Two-class classification (HS vs. LBP) accuracy results (with ten-fold cross-validation) with secondary features sets created based on each single primary feature (Table 3.2) only, from left to right, across all classifiers. Classifiers employed are decision trees (DT), discriminant analysis (LDA and QDA), logistic regression (LR), support vector machines (SVM), nearest neighbor (*k*NN), and ensemble methods (bagged, boosted, subspace, random undersampling - RUS), from top to bottom. For each classifier, the primary feature providing the highest accuracy is shown in boldface. Median (Median) and maximum (Max) values for each individual primary feature across all classifiers are shown as well.

Classification model	ZC	SSC	WAMP	MAV	IEMG	VAR	RMS	WL	LD	KURT	SKEW	PE	MDF	RVD
Fine DT	0.69	**0.78**	0.73	0.72	0.69	0.75	0.73	0.63	0.69	0.67	0.64	0.71	0.70	0.67
Medium DT	0.72	**0.79**	0.73	0.72	0.68	0.76	0.75	0.61	0.70	0.68	0.66	0.72	0.70	0.69
Coarse DT	0.67	**0.78**	0.72	0.70	0.70	0.77	0.76	0.70	0.69	0.63	0.66	0.70	0.70	0.67
LDA	0.76	0.79	0.77	0.76	0.76	0.78	0.75	0.75	0.76	0.71	0.75	**0.81**	0.71	0.78
QDA	0.78	**0.79**	0.74	0.69	0.69	0.79	0.68	0.75	0.78	0.72	0.68	**0.79**	0.68	0.73
LR	0.75	**0.79**	0.76	0.76	0.76	0.77	0.75	0.75	0.77	0.70	0.74	0.78	0.71	**0.79**
Linear SVM	0.75	0.78	0.77	0.74	0.73	0.77	0.77	0.75	0.75	0.73	0.75	**0.83**	0.71	0.78
Quadratic SVM	0.78	0.84	0.85	0.71	0.72	0.76	0.76	0.70	0.78	0.73	0.72	**0.86**	0.70	0.79
Cubic SVM	0.82	0.77	0.76	0.67	0.69	0.79	0.74	0.60	0.78	0.68	0.63	**0.83**	0.77	0.74
Fine Gaussian SVM	**0.74**	0.66	0.66	0.59	0.59	0.71	0.58	0.65	0.59	0.59	0.71	0.63	0.62	0.66
Medium Gaussian SVM	0.81	0.84	0.84	0.71	0.72	0.75	0.80	0.73	0.79	0.78	0.72	**0.86**	0.72	0.79
Coarse Gaussian SVM	0.79	0.80	0.77	0.67	0.68	0.76	0.77	0.76	0.71	0.70	0.75	**0.82**	0.73	0.78
Fine kNN	0.74	**0.88**	0.84	0.72	0.73	0.75	0.74	0.67	0.66	0.74	0.62	0.79	0.78	0.80
Medium kNN	0.77	0.81	0.81	0.72	0.72	0.78	0.78	0.74	0.74	0.75	0.73	**0.85**	0.73	0.74
Coarse kNN	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59
Cosine kNN	0.77	0.80	0.82	0.73	0.73	0.75	0.73	0.77	0.75	0.73	0.72	**0.83**	0.74	0.80
Cubic kNN	0.77	0.81	**0.82**	0.72	0.71	0.77	0.76	0.74	0.75	0.75	0.74	0.81	0.72	0.73
Weighted kNN	0.76	0.84	**0.86**	0.74	0.74	0.78	0.78	0.71	0.73	0.75	0.69	**0.86**	0.78	0.82
Boosted Trees[a]	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.43	0.41	0.41	0.52	0.41	0.41	0.41
Bagged Trees	0.75	**0.81**	0.79	0.72	0.71	0.76	0.76	0.70	0.77	0.72	0.69	0.79	0.77	0.76
Subspace Discriminant	0.77	0.81	0.75	0.73	0.74	0.76	0.77	0.78	0.73	0.73	0.76	**0.82**	0.73	0.77
Subspace kNN	0.75	**0.83**	0.82	0.72	0.72	0.73	0.75	0.71	0.70	0.77	0.67	0.82	0.78	0.78
RUS Boosted	0.71	**0.79**	0.74	0.70	0.71	0.73	0.72	0.65	0.76	0.68	0.66	0.75	0.74	0.69
Median	0.75	0.79	0.77	0.72	0.71	0.76	0.75	0.71	0.74	0.72	0.69	**0.81**	0.72	0.76
Max	0.82	**0.88**	0.86	0.76	0.76	0.79	0.8	0.78	0.79	0.78	0.76	0.86	0.78	0.82

[a]Underperforming classifier(s) with classification accuracy <50%, for most of the feature sets.

classification accuracy of 0.88 for *k*NN classifier with fine resolution ($k = 1$). These results were reflected in the calculated average ranks where SSC and PE were labeled as the best and second best, respectively.

Taking into account the classification results, respective assigned ranks, as well as correlation insights obtained in section 5.1.1, an initial subset of five primary features was selected: SSC, WAMP, VAR, ZC, RVD. Features that were demonstrating high correlation (MAV, IEMG, LD, RMS), or limited classification contribution (MDF, WL, KURT, SKEW), were removed from the set. The contribution of PE was analyzed and consequently was added to the selected set, despite exhibiting a high correlation with SSC. PE demonstrated the highest median values (0.81) across all classifiers, compared to 0.79 for SSC, and the 2nd best average rank result (Ta-

Table 5.2: Sorted average rank list where features with the best average classification rank across all classifiers are set on top.

Primary feature	Average rank
SSC	2.38
PE	2.62
WAMP	4.10
VAR	5.05
ZC	6.17
RVD	6.29
RMS	7.40
LD	8.19
MDF	9.45
WL	10.07
MAV	10.45
IEMG	10.57
KURT	10.71
SKEW	11.55

ble 5.2). Therefore, PE could not have been ignored despite its correlation with the best-ranked SSC feature.

5.3.3 Contextual feature sets confirmation

Altogether insights from the previous steps were exploited to create several concurrent secondary feature sets. These sets were employed to compare and validate the classification outcomes for the given data set, consisting of HS and LBP (CLBP and RLBP) groups. A difference among secondary feature sets was in the list of primary features utilized as the input for subsequent construction of all secondary feature groups. Thus, the raw primary features employed, were as follows:

1.Complete primary features set with no NCA feature selection performed
2.Complete primary features set with NCA feature selection performed
3.Hudgin's vector (MAV, WL, ZC, SSC) [123]
4.Du's vector (IEMG, VAR, WL, ZC, SSC, WAMP) [124]
5.Top five selected primary features as per section 5.3.2 (SSC, WAMP, VAR, ZC, RVD), labelled as "Top 5"
6.Top five selected primary features as per section 5.3.2 with additionally including PE (SSC, WAMP, VAR, ZC, RVD + PE), labelled as "Top 5+PE"
7.Nine best uncorrelated primary features as per section 5.3.2 (ZC, WAMP, VAR, WL, KURT, SKEW, PE, MDF, RVD) with PE instead of its correlated counterpart SSC, labeled

as "Uncorrelated"

Classification results for the lists of the selected primary features sets are presented in Table 5.3. It can be observed that the "Uncorrelated" feature set exhibited the best median classification accuracy (0.85), alongside providing the best maximum classification accuracy of 0.95 for fine and subspace *k*NN classifier, but at the expense of nine primary features employed and twenty-nine NCA components selected. The "Top 5+PE" set provided similarly good results, with the same number of best classification results per classifier (seven) and with a smaller number of NCA components ($N_{NCA} = 20$). The worst performing feature set was the one employing all features ($N_{Tot} = 727$), without NCA feature selection. Feature set "Top 5" resulted in the least number of NCA components ($N_{NCA} = 14$).

Further, one of the main goals of this step was to examine the contribution of each secondary feature group and enable inference on the contribution of such feature set to the classification results. Therefore, a separate classification procedure was performed where each individual secondary feature group (Coordination, Co-activation, Trends, Fatigue indices as per section 4.5) was chosen as an input into the classifier models (again, the same twenty-three classifier variants). The intention was to verify how well each individual secondary set deals with a given detection of LBP patients and differentiations tasks and consequently to enable inference about the potential of such feature construction and justification behind. Top 5+PE were selected (SSC, WAMP, VAR, ZC, RVD + PE) as an underlying subset for constructing each secondary group.

Results are presented in Table 5.4. It can be observed that the Coordination features group exhibited overall best accuracy, among secondary feature groups, with a median value of 0.81 across all classifiers, closely followed by the Co-activation features group with 0.80 median value. Trends, Fatigue indices, and Primary groups exhibited notably less successfully, with the Primary group exhibiting even somewhat better than the remaining two groups. Overall best accuracy result was achieved for the Complete set (all secondary groups combined), with 0.85 median value. Additionally, it can be noticed that Coordination groups performed significantly better for decision trees type of classifiers (accuracy ranging from 0.82 to 0.85) compared to other feature sets, whereas the Complete set significantly outperformed the SVM classifiers with accuracy results >0.90.

5.3.4 Comparison among subjects groups

One of the main tasks was to provide insights into the proposed method's capability of dealing with differentiation among multiple (more homogeneous) subjects groups, thus going a step beyond only separating the healthy subject (HS) from patients (LBP). Therefore, another classification iteration was conducted with an aim to determine groups' differentiation success for (I) HS vs. LBP, (II) HS vs. RLBP, (III) HS vs. CLBP, and (IV) CLBP vs. RLBP pairs. Patients'

Table 5.3: Two-class classification (HS vs. LBP) accuracy results (with ten-fold cross-validation) for concurrent, primary features-based, secondary feature sets, from left (1) to right (7), as defined in section 5.3.3. Feature sets (2 to 7) are employing NCA feature selection with the number of components given by N_{NCA}. For feature set (1), all features were employed (N_{Tot}). The best performing feature set for the given classifier is shown in boldface. Classifiers with accuracy results $\geq$90% are shown in boldface. Median (Median) and maximum (Max) values for each secondary feature set across all classifiers are shown as well.

Classification model	(1) All (no NCA) $N_{Tot} = 727$	(2) All (with NCA) $N_{NCA} = 31$	(3) Hudgin $N_{NCA} = 16$	(4) Du $N_{NCA} = 20$	(5) Top 5 $N_{NCA} = 14$	(6) Top 5+PE $N_{NCA} = 20$	(7) Uncorrelated $N_{NCA} = 29$
Fine DT	0.75	0.77	0.82	0.78	**0.83**	0.80	0.79
Medium DT	0.75	0.78	0.81	0.78	**0.84**	0.78	0.80
Coarse DT	0.75	0.79	0.76	0.78	**0.80**	0.76	0.75
LDA	0.71	0.75	0.78	0.73	0.81	**0.84**	0.82
QDA[a]	N/A	N/A	N/A	N/A	**0.85**	0.75	N/A
LR	0.58	0.74	0.80	0.75	0.80	**0.84**	0.80
Linear SVM	0.78	0.77	0.83	0.79	0.81	**0.86**	0.85
Quadratic SVM	0.78	0.81	0.84	0.82	0.83	**0.94***	0.92*
Cubic SVM	0.80	0.86	0.84	0.81	0.82	**0.94***	0.92*
Fine Gaussian SVM	0.59	0.59	0.59	0.59	0.59	0.59	0.59
Medium Gaussian SVM	0.78	0.87	0.87	0.86	0.85	**0.93***	0.92*
Coarse Gaussian SVM	0.70	0.76	0.83	0.68	0.79	0.83	0.79
Fine kNN	0.72	0.92*	0.88	0.82	0.91*	0.90*	**0.95***
Medium kNN	0.75	0.79	0.82	0.80	0.85	0.84	**0.87**
Coarse kNN	0.59	0.59	0.59	0.59	0.59	0.59	0.59
Cosine kNN	0.76	0.82	0.83	0.78	0.82	0.82	**0.89**
Cubic kNN	0.75	0.81	0.83	0.80	0.83	0.86	**0.87**
Weighted kNN	0.74	0.82	0.84	0.81	0.87	0.85	**0.90***
Boosted Trees[b]	0.41	0.41	0.41	0.41	0.41	0.41	0.41
Bagged Trees	0.76	0.78	0.83	0.85	**0.85**	0.84	**0.85**
Subspace Discriminant	0.76	0.82	0.81	0.83	0.84	**0.86**	0.85
Subspace kNN	0.71	0.93*	0.87	0.87	0.89	0.90*	**0.95***
RUS Boosted	0.76	0.76	0.82	0.77	**0.85**	0.78	0.80
Median	0.75	0.79	0.82	0.79	0.83	0.84	**0.85**
Max	0.80	0.93	0.88	0.87	0.91	0.94	**0.95**

*Classification results with accuracy $\geq$90%.

[a]Some of QDA classification models failed due to a singular covariance matrix for one of the classes.

[b]Underperforming classifier(s) with classification accuracy <50%.

groups, namely LBP, RLBP, and CLBP, in comparison with HS, were labeled as a positive class (1), whereas HS was labeled as a negative class (0). In comparison between CLBP and RLBP, RLBP was labeled a positive class (1). ACC, PPV and RC as measures of classification success were also calculated and given in Table 5.5.

Overall, noticeably, the best classification results were achieved for HS vs. RLBP differentiation with classification accuracy ranging from 0.93 to 0.98 for all classifiers except coarse *k*NN, boosted trees and RUS boosted with the accuracy of 0.60, 0.60, and 0.78, respectively. Moreover, for HS vs. RLBP differentiation, sensitivity across all classifiers was in the range

Table 5.4: Two-class classification (HS vs. LBP) accuracy results (with ten-fold cross-validation) for different secondary feature groups (chapter 4). The top 5+PE subset of primary features is employed. Primary features (Primary max), leftmost group, together with a complete feature set consisting of all secondary groups (Complete set), rightmost, are validated and shown as well. NCA feature selection procedure was performed for all feature groups with the number of components given by N_{NCA}. The best performing feature group for the given classifier is shown in boldface. Classifiers with any of the accuracy results ≥90% are shown in boldface. Median (Median) and maximum (Max) values for each secondary feature group across all classifiers are shown as well.

Classification model	Primary max $N_{NCA}=6$	Fatigue indices $N_{NCA}=7$	Trends $N_{NCA}=7$	Coordination $N_{NCA}=11$	Co-activation $N_{NCA}=12$	Complete set $N_{NCA}=20$
Fine DT	0.66	0.63	0.72	**0.84**	0.68	0.80
Medium DT	0.69	0.65	0.71	**0.85**	0.68	0.78
Coarse DT	0.64	0.56	0.72	**0.82**	0.68	0.76
LDA	0.77	0.58	0.76	0.78	0.82	**0.84**
QDA[a]	N/A	**0.80**	0.76	**0.80**	0.75	0.75
LR	0.76	0.60	0.77	0.80	0.80	**0.84**
Linear SVM	0.76	0.62	0.77	0.80	0.84	**0.86**
Quadratic SVM	0.79	0.65	0.81	0.82	0.80	**0.94***
Cubic SVM	0.78	0.61	0.79	0.82	0.88	**0.94***
Fine Gaussian SVM	0.61	**0.75**	0.70	0.67	0.59	0.59
Medium Gaussian SVM	0.78	0.63	0.78	0.82	0.84	**0.93***
Coarse Gaussian SVM	0.74	0.59	0.75	0.77	0.77	**0.83**
Fine kNN	0.79	0.66	0.73	**0.91***	0.85	0.90*
Medium kNN	0.80	0.58	0.79	0.81	0.81	**0.84**
Coarse kNN	0.59	0.59	0.59	0.59	0.59	0.59
Cosine kNN	0.79	0.61	0.79	0.81	0.77	**0.82**
Cubic kNN	0.81	0.65	0.80	0.81	0.80	**0.86**
Weighted kNN	0.79	0.70	0.79	0.84	0.82	**0.85**
Boosted Trees[b]	0.41	0.43	0.41	0.41	0.41	0.41
Bagged Trees	0.79	0.66	0.79	0.80	0.80	**0.84**
Subspace Discriminant	0.77	0.55	0.76	0.79	0.81	**0.86**
Subspace kNN	0.71	0.63	0.78	**0.90***	0.84	**0.90***
RUS Boosted	0.75	0.63	0.74	**0.80**	0.72	0.78
Median	0.77	0.63	0.76	0.81	0.80	**0.84**
Max	0.81	0.80	0.81	0.91	0.88	**0.94**

*Classification results with accuracy ≥90%.

[a]Some of QDA classification models failed due to a singular covariance matrix for one of the classes.

[b]Underperforming classifier(s) with classification accuracy <50%.

0.95 to 1.00, where precision for all classifiers, except coarse *k*NN, boosted trees and RUS boosted, was in the range 0.90 to 0.99.

The second-best classification results were obtained for HS vs. LBP classification with the best accuracy achieved for SVM variants, namely, for quadratic, cubic and medium Gaussian kernel classifiers, with 0.93, 0.94, 0.94, respectively. Sensitivity results >0.95 were achieved for quadratic and cubic SVM, 0.96 and 0.97, respectively. The highest result for sensitivity was obtained with boosted trees (100%), but, similarly as for fine Gaussian SVM and coarse *k*NN,

Table 5.5: Two-class classifier (with ten-fold cross-validation) results for accuracy (ACC), precision (PPV), sensitivity (RC) are shown. A series of classification differentiation pairs are (I) HS vs. LBP, (II) HS vs. RLBP, (III) HS vs. CLBP, (IV) CLBP vs. RLBP for healthy (HS) subjects, subjects with radiculopathy (RLBP), chronic low back pain patients (CLBP), and LBP as a joint group for RLBP and CLBP patients. ACC, PPC, and RC results ≥90% are shown in boldface. Highlighted rows show the best-performing classifiers. Median (Median) and maximum (Max) values for each classification pair for ACC, PPV, and RC are calculated.

Classification model	ACC				PPV				RC			
	(I)	(II)	(III)	(IV)	(I)	(II)	(III)	(IV)	(I)	(II)	(III)	(IV)
Fine DT	0.77	**0.94**	0.74	0.68	0.72	**0.94**	0.76	0.71	0.74	**0.95**	0.79	0.68
Medium DT	0.78	**0.94**	0.73	0.69	0.73	**0.94**	0.75	0.72	0.74	**0.96**	0.79	0.70
Coarse DT	0.73	**0.93**	0.69	0.69	0.69	**0.94**	0.72	0.71	0.63	**0.95**	0.75	0.70
LDA	0.85	**0.98**	0.75	0.74	0.77	**0.97**	0.77	0.78	0.88	**1.00**	0.78	0.73
QDA[a]	0.75	**0.96**	0.78	N/A[a]	0.78	**0.97**	0.80	N/A[a]	0.52	**0.96**	0.80	N/A[a]
LR	0.83	**0.95**	0.75	0.75	0.76	**0.95**	0.77	0.77	0.86	**0.97**	0.78	0.76
Linear SVM	0.87	**0.96**	0.76	0.67	0.81	**0.97**	0.77	0.71	0.89	**0.97**	0.81	0.67
Quadratic SVM	**0.93**	**0.97**	0.84	0.83	0.89	**0.97**	0.86	0.83	**0.96**	**0.98**	0.86	0.85
Cubic SVM	**0.94**	**0.97**	0.88	0.82	0.89	**0.97**	0.89	0.84	**0.97**	**0.98**	**0.90**	0.82
Fine Gaussian SVM[b]	0.59	**0.93**	0.56	0.54	N/A[b]	**0.90**	0.56	0.54	0.00[b]	**1.00**	**1.00**	**1.00**
Medium Gaussian SVM	**0.94**	**0.98**	0.85	0.77	**0.93**	**0.97**	0.85	0.78	**0.91**	**0.99**	0.89	0.81
Coarse Gaussian SVM	0.82	**0.98**	0.74	0.64	0.82	**0.97**	0.69	0.60	0.72	**1.00**	**0.99**	**0.96**
Fine kNN	**0.90**	**0.98**	**0.90**	0.89	0.85	**0.97**	**0.91**	**0.92**	**0.92**	**1.00**	**0.92**	0.87
Medium kNN	0.84	**0.98**	0.82	0.74	0.75	**0.97**	0.78	0.73	**0.90**	**1.00**	**0.94**	0.81
Coarse kNN[b]	0.59	0.60	0.56	0.54	N/A[b]	0.60	0.56	0.54	0.00[b]	**1.00**	**1.00**	**1.00**
Cosine kNN	0.83	**0.98**	0.82	0.74	0.74	**0.97**	0.79	0.75	0.88	**1.00**	**0.92**	0.79
Cubic kNN	0.85	**0.98**	0.78	0.74	0.78	**0.97**	0.76	0.73	**0.90**	**1.00**	**0.91**	0.80
Weighted kNN	0.85	**0.98**	0.80	0.80	0.77	**0.97**	0.77	0.79	**0.90**	**1.00**	**0.92**	0.84
Boosted Trees	0.41	0.60	0.56	0.54	0.41	0.60	0.56	0.54	**1.00**	**1.00**	**1.00**	**1.00**
Bagged Trees	0.86	**0.95**	0.79	0.74	0.82	**0.95**	0.77	0.74	0.85	**0.98**	0.88	0.79
Subspace Discriminant	0.85	**0.97**	0.77	0.74	0.80	**0.95**	0.76	0.76	0.84	**1.00**	0.85	0.77
Subspace kNN	**0.90**	**0.96**	**0.90**	0.86	0.84	**0.99**	0.87	0.85	**0.94**	**0.95**	**0.97**	**0.90**
RUS Boosted	0.79	0.78	0.74	0.69	0.71	0.74	0.75	0.71	0.80	**0.98**	0.81	0.73
Median	0.84	0.96	0.77	0.74	0.78	0.97	0.77	0.74	0.88	0.99	0.89	0.81
Max	0.94	0.98	0.90	0.89	0.93	0.99	0.91	0.92	1.00	1.00	1.00	1.00

[a]Some of QDA classification models failed due to a singular covariance matrix for one of the classes.

[b]Classification failed due to all subjects being classified to only one (negative) class label.

resulted in 0% sensitivity and undefined precision. This was due to having all samples labeled with either positive or negative class, thus these results were considered fallacious.

For HS vs. CLBP differentiation, the best accuracy results of 0.90 were obtained by fine resolution *k*NN and subspace *k*NN. Fine *k*NN was also the only classifier exhibiting precision >0.90, whereas for sensitivity, multiple classifiers exhibited values >0.90, with fine Gaussian SVM, coarse *k*NN and boosted trees moreover exhibiting sensitivity of 100%, but at the expense of low accuracy and precision. Cubic SVM, fine *k*NN, and subspace *k*NN provided a good balance of classification success for HS vs. CLBP across all measured values (ACC, PPV, RC).

The least performing classification differentiation pair was for CLBP vs. RLBP. None of the classifiers achieved accuracy >0.90, where the best classification accuracy of 0.89 was achieved for fine *k*NN. The same classifier was the only one to exhibit a precision >0.90, namely 0.92. Apart from the low-performing classifiers (fine Gaussian SVM, coarse *k*NN, and boosted trees), the best performing sensitivity was demonstrated by fine *k*NN and subspace *k*NN, with 0.87 and 0.90, respectively.

Chapter 6

Clustering and CLBP patients subgrouping

In this chapter, the focus is put on providing insights into the possibility of patients' subgrouping based on the subset of features selected as part of the proposed LBP contextual features modeling. The main prerequisite was to keep features in the original domain, which was achieved utilizing the NCA procedure (section 5.1.2). Following the classification results analyzed (Table 5.5), it was confirmed that the selected features have very good discriminatory power in terms of separating healthy subjects from the LBP patients. Moreover, good results in further differentiating healthy subjects from CLBP, as well as a differentiation between CLBP and RLBP groups, were observed.

Subsequently, the main attention in this chapter is directed toward finding more specific, homogeneous, subgroups within the CLBP group itself. This was achieved by creating CLBP profiles for each CLBP subject. From this, the initial set of subgroups is proposed. Subsequently, by employing more formal clustering techniques, these groups were confirmed and expanded with additional grouping clusters, resulting in seven CLBP subgroups total. As part of the subgroups detection and assignment procedure, additional metrics are introduced to enable an objective procedure for CLBP patients subgrouping.

6.1 Differentiation results

As already mentioned, one of the key outcomes of the so far presented method is the ability to preserve the constructed contextual features in the original domain. Individual two-class classification results were accompanied by the respective feature sets based on NCA selection, as presented in Table 6.1.

It can be noticed that each classification pair has its own specific set of NCA features selected, with a different number of NCA components assigned. Namely, twenty, three, thirteen,

and fourteen for HS vs. LBP, HS vs. RLBP, HS vs. CLBP, and CLBP vs. RLBP, respectively. This comes along with the observations and statements about the complex neuromotor processes that need to be modeled for particular differentiation and subgrouping cases, leading eventually to more individualized approaches.

Differentiation between HS and LBP groups posed certain complexity due to three different groups analyzed in behind (namely, the LBP group was combined from CLBP and RLBP patients), thus the model had to account for certain diversity in behind, which resulted in twenty NCA components selected and 4 out of 5 secondary feature groups included in the models. Nonetheless, classification accuracy results of up to 0.94 and sensitivity of up to 0.97 (Table 5.5) provided a clear notion of good discriminative power of the proposed feature modeling approach.

On the other hand, HS vs. RLBP differentiation was shown to be the most distinctive one with only three NCA components involved and having in mind high classification accuracy results (>0.96), it was evident that very clear and distinctive differences between these two groups were captured with the presented modeling. This nice separation among HS and RLBP groups is visualized in Fig. 6.1.

However, the main goal of this PhD thesis was to enable more insights into the non-specific CLBP group, consequently leading to CLBP patients subgrouping, along with providing meaningful medically-based interpretation. Namely, the non-specific CLBP patients subgrouping has been clearly emphasized on multiple occasions to be the main prerequisite for successful LBP diagnostics and rehabilitation treatments [7, 8, 25]. Differentiation between CLBP and HS groups posed certain challenges as it was evident from classification accuracy results in the range 0.88 to 0.90, which was somewhat less compared to LBP vs. HS differentiation with accuracy results >0.90. Nevertheless, the obtained results are significantly high to provide confidence in the selection of the proposed models, especially taking into consideration the challenging tasks at hand.

Insightfully, the most challenging classification task was between CLBP and RLBP groups, with classification accuracy up to 0.89 and fourteen NCA components employed across all five secondary features groups (Coordination, Co-activation, Trends, Fatigue indices, and Primary features). The characteristics of RLBP that led to significantly high discrimination results, as in HS and RLBP pair, or those that contributed to very good classification results, as for HS vs. LBP task, were evidently not so discriminative within a group of subjects with LBP pathology itself, namely within LBP group (CLBP and RLBP), for a CLBP vs. RLBP differentiation task. This again aligns with the considerations of differentiating and subgrouping LBP (or more precisely, CLBP) patients being the most challenging task.

Table 6.1: NCA components lists and comparison as a result of different subject groups classification and differentiation for HS vs. LBP, HS vs. RLBP, HS vs. CLBP, CLBP vs. RLBP, from left to right, respectively. Secondary feature group (Group), feature metric (Metric), and involved relations among muscle locations (Relation), as feature insights, are listed for each classification group. Groups labeled CO, CA, TR, FI, and PR, stand for Coordination, Co-activation, Trends, Fatigue indices, and Primary-derived secondary feature groups, respectively. For CA group, s- and u- stand for respective co-activation alignment ("sync") and misalignment ("unsync") scores. Muscle location relations are expressed in terms of different relation combinations for ULES (UL) and LLES (LL) lumbar regions, left (L) and right side, and starting (st.) or ending (end) segment of the exercise. The number of NCA components are twenty, three, thirteen, and fourteen for HS vs. LBP, HS vs. RLBP, HS vs. CLBP, and CLBP vs. RLBP, respectively.

NCA	HS vs. LBP			HS vs. RLBP			HS vs. CLBP			CLBP vs. RLBP		
	Group	Metric	Relation	Group	Metric	Relation	Group	Metric	Relation	Group	Metric	Relation
1	PR	max ZC	LL-R	TR	Tr-2 SSC	min UL	TR	Tr-1 SSC	min LL	PR	max ZC	LL-R
2	PR	max SSC	UL-L	CO	Corr SSC	LL L-R	TR	Tr-2 SSC	min UL	TR	Tr-4 VAR	LL L-R st.
3	PR	max VAR	LL-R	CA	s-RVD	LL-L	TR	Tr-4 WAMP	LL L-R end	TR	Tr-3 RVD	max LL-UL
4	TR	Tr-3 ZC	max LL-UL				TR	Tr-4 RVD	LL L-R end	TR	Tr-4 PE	LL L-R st.
5	TR	Tr-1 SSC	min LL				TR	Tr-1 PE	min LL	FI	FI-0	LL-R f0
6	TR	Tr-2 SSC	min UL				CO	Corr VAR	LL L-R	FI	FI-1	LL L-R k
7	TR	Tr-3 SSC	min LL-UL				CO	Corr PE	LL L-R	FI	FI-6	LL-R
8	TR	Tr-4 RVD	LL L-R st.				CA	s-SSC	UL-R	CO*	DTW SSC	LL L-R
9	TR	Tr-4 RVD	LL L-R end				CA	s-WAMP	LL-L	CO*	DTW RVD	UL L-R
10	CO	Corr VAR	LL L-R				CA	u-WAMP	LL-L	CO*	Spear PE	UL L-R
11	CO	Corr RVD	UL L-R				CA*	s-SSC	UL-L	CA	s-PE	LL-R
12	CA	s-VAR	LL-R				CA*	u-WAMP	LL-L	CA*	u-WAMP	LL-L
13	CA	u-VAR	UL-R				CA*	s-PE	UL-R	CA*	s-RVD	UL-L
14	CA	s-WAMP	LL-R							CA*	u-RVD	UL-R
15	CA	u-WAMP	UL-L									
16	CA	s-RVD	LL-L									
17	CA	s-RVD	UL-R									
18	CA*	s-WAMP	LL-L									
19	CA*	s-RVD	UL-R									
20	CA*	s-PE	UL-R									

*Groups with auto-correlation derived feature.

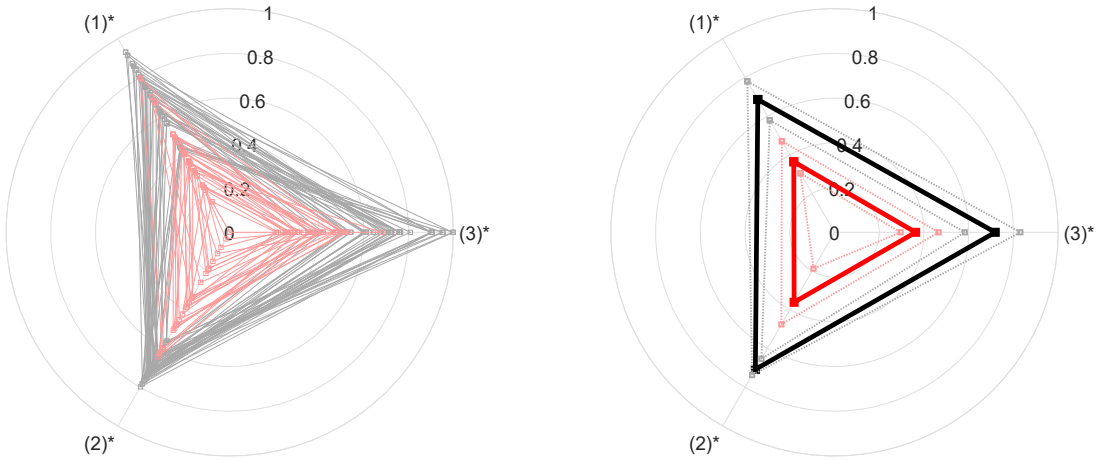


Figure 6.1: NCA component values are shown in spider plots for comparison between HS and RLBP subjects. NCA component values for each HS (gray) and RLBP (red) subject, are shown on left. Median values for each feature component, for HS and RLBP group, in black and red solid contours, shown on right, respectively. Dotted contours, inner and outer, for 25- and 75-percentile, shown for HS and RLBP group, in gray and light red, respectively. Three resulting NCA components, (1) to (3), as listed in Table 6.1 under HS vs. RLBP comparison. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

6.2 Differentiation models inspection

As the next step, a deeper inspection of the differentiation classification results was undertaken. Concretely, NCA features obtained (Table 6.1) were further examined to verify the possibility of interpreting results with the relevant medical-based background. CLBP patients group was of primary interest, therefore feature models for the respective CLBP patients were examined in more detail. To enable this deeper insight, for each individual CLBP patient a respective CLBP profile was created. This was achieved by introducing radial (spider) plots, thus giving a visual representation of myoelectric-related LBP characteristics for each CLBP patient.

CLBP profile prototypes were constructed as presented in Fig. 6.2. For each patient, a respective individual's CLBP profile was drawn as a contour (blue) stretched by thirteen NCA components in radial representation. NCA components were normalized in the range 0 to 1, across the whole NCA samples set for the respective HS vs. CLBP classification pair. Each profile is accompanied by the median-based calculated profile for all CLBP (red contours) and HS (black contours) subjects, thus providing the visual indication of how well the respective individual profile inclines to or deviates from the "averaged" CLBP or HS profiles. Further, to make the respective profile more intuitive, NCA components were organized in such a manner to represent the muscle sites topology (geometric placement) in the following way: components (8, 13, 11, 2) relating to the ULES region were placed on top of the plot, and the remaining components relating to the LLES region were placed on the bottom. Also, the notion of the respective left and right side of muscle sites was followed where possible.

Additionally, to bring more insight into the differences between CLBP and HS feature val-

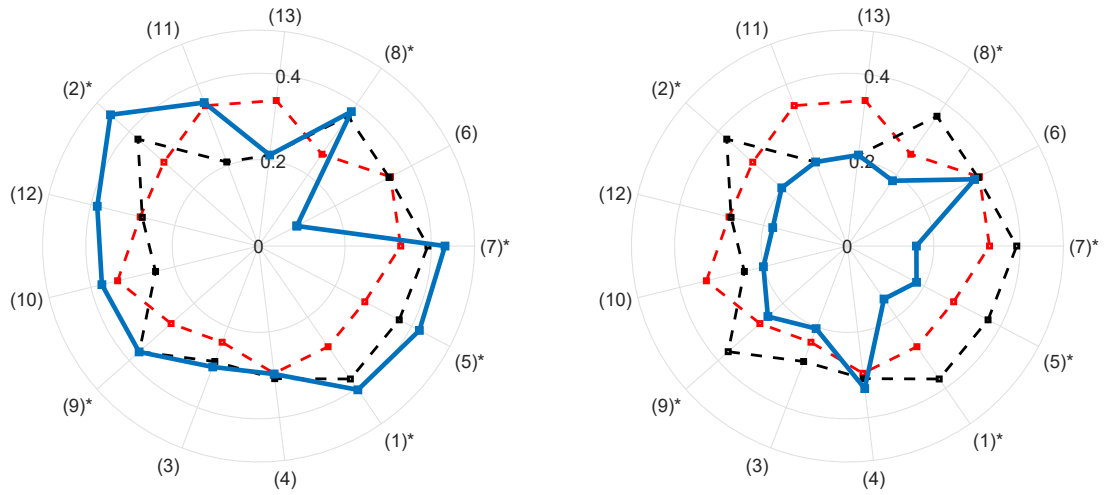


Figure 6.2: NCA component values for two individuals with CLBP are shown in a spider plot with blue contours (left and right). Median component values across all HS and CLBP subjects are shown in black and red dashed contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.1. Components are organized in such a manner to follow the geometric placement of muscle sites, with components 8, 13, 11, and 2 relating to ULES (top) and the remaining components relating to LLES (bottom). The notion for left and right geometric placement is followed where possible. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

ues, for each NCA component Wilcoxon rank sum test at 5% significance level was calculated. The Wilcoxon rank sum test is a nonparametric test (equivalent to the Mann-Whitney U-test) for equality of population medians of two independent samples X and Y [140, 144]. Components annotated with (*) were detected to have median values differing between CLBP and HS with statistical significance. These components (with order numbers assigned as per Table 6.1) were as follows: 1, 2, 5, 7, 8, and 9, placed both in the ULES and LLES regions and found across all feature group in this HS vs. CLBP classification pair (Coordination, Co-activation, and Trends).

To further analyze the features and factors that are contributing to CLBP differentiation, each NCA component (for HS vs. CLBP pair) was inspected more thoroughly, alongside providing the possible explanation for its contribution to the respective classification task. These insights are summarized in a Table 6.2. For each component, the respective Wilcoxon rank sum test p-value, type of information carried by respective feature, and explanation are presented, alongside earlier presented Group, Metric, and Relation properties as in Table 6.1. It can be observed that the majority of components carry some type of frequency information, where some are related to trends and assumably fatiguing processes, whereas some are related to coordination and (co-)activation patterns potentially pointing to different types of motor units recruited during the exercise duration, differing in number, dominant muscle fiber type (e.g., Type I slow-twitch or Type II fast-twitch fibers) or motor unit size itself. Though, it is worth noticing that for other classification pair tasks (Table 6.1), the frequency-related type of information was less dominant, i.e. components carrying non-linear or energy complexity information where more

present, as in e.g. CLBP vs. RLBP differentiation pair.

In the context of the aforementioned frequency-related features dominance, for HS vs. CLBP pair, further inspection and analysis in this PhD thesis can provide more insights into whether this interpretation (as presented in Table 6.2) can be related to other research works showing that CLBP subjects have a higher presence of fast-twitch type II muscle fibers in the back muscles [51, 169]. Namely, as per prototype profiles presented in Fig. 6.2, it is evident that CLBP patients' profiles can differ quite significantly, thus implying completely different neuromotor control mechanisms behind, extending to possibly different physiological processes or histological muscle fibers compositions as well which do not necessarily follow one uniform pattern. This again concords with the inaugurated view that the "one-size fits all" approach is not suitable for deeper inference on this problematic [8, 25].

Table 6.2: NCA features list with thirteen components (ID column) resulting from HS vs. CLBP classification differentiation task (Table 6.1). Components' Wilcoxon p-value, Group, Metric, Relation, Type of information, and Explanation are given, from left to right column-wise, respectively.

NCA ID	Wilcoxon p-value	Group	Metric	Relation	Type of information	Feature explanation
1	p=0.0004*	TR	Tr-1 SSC	min LL	Frequency	Points to a significant difference between CLBP and HS for the LLES side that is more supporting higher frequency activation triggers.
2	p=0.0157*	TR	Tr-2 SSC	min UL	Frequency	Points to a significant difference between CLBP and HS for the ULES side that is more supporting higher frequency activation triggers.
3	p=0.1609	TR	Tr-4 WAMP	LL L-R end	Frequency, activation	Indicates differences in motor unit action potential triggers and muscle force between the left and right sides at the end of contraction.
4	p=0.1443	TR	Tr-4 RVD	LL L-R end	Activation, complexity	Points to the different patterns of muscle force levels or activation bursts between LLES left and right sides at the end of the exercise.
5	p=0.0010*	TR	Tr-1 PE	min LL	Non-linearity, complexity	Points to a statistically significant difference between CLBP and HS for the LLES side that is supporting more complex or irregular muscle activity or neurocoding patterns.
6	p=0.2246	CO	Corr VAR	LL L-R	Energy, force	Points to differences in maintaining coordinated exerted force levels between left and right LLES sides, with comparison between CLBP and HS.
7	p=0.0022*	CO	Corr PE	LL L-R	Non-linearity, complexity	Points to differences in maintaining coordinated patterns of muscle activity complexity between left and right LLES sides, assumably related to neuromotor control activation strategies of different motor units and fiber types, with comparison between CLBP and HS.
8	p=0.0005*	CA	s-SSC	UL-R	Frequency	The higher the value (score), the better synchronization of higher frequency activation triggers between the right side of ULES and the rest of the muscles sites in ULES and LLES region, which is with statistical significance more pertained to HS, compared to CLBP.
9	p=0.0002*	CA	s-WAMP	LL-L	Frequency, activation	The higher the value (score), the better synchronization of higher activity triggering and muscle force between the left side of LLES and the rest of the muscles sites in ULES and LLES region, which is with statistical significance more pertained to HS, compared to CLBP.
10	p=0.0683	CA	u-WAMP	LL-L	Frequency, activation	The higher the value (score), the worse synchronization of higher activity triggering and muscle force between the left side of LLES and the rest of the muscles sites in ULES and LLES region, which is more pertained to CLBP, but without clear statistical significance.
11	p=0.1601	CA[1]	s-SSC	UL-L	Frequency	Processed with autocorrelation, thus enabling detecting and emphasizing more subtle synchronization patterns among muscle sites; the higher the value (score), the better synchronization of higher frequency of activation triggers between the left side of ULES and the rest of the muscles sites in ULES and LLES region, which is more pertained to CLBP, but without clear statistical significance.
12	p=0.3881	CA[1]	u-WAMP	LL-L	Frequency, activation	Processed with autocorrelation, thus enabling detecting and emphasizing more subtle synchronization patterns among muscle sites; the higher the value (score), the worse synchronization of higher activity triggering and muscle force between the left side of LLES and the rest of the muscles sites in ULES and LLES region, with results showing to equally pertain to both CLBP and HS in general.
13	p=0.3068	CA[1]	s-PE	UL-R	Non-linearity, complexity	Processed with autocorrelation, thus enabling detect ing and emphasizing more subtle synchronization patterns among muscle sites; the higher the value (score), the better synchronization of more complex muscle activity patterns between the right side of ULES and the rest of the muscle sites in the ULES and LLES region, and more pertained to CLBP compared to HS.

*Wilcoxon rank sum test null hypothesis rejected at 5% significance level.

[1]Groups with auto-correlation derived feature.

6.3 CLBP patients profiles

As previously shown by CLBP profile prototypes (Fig. 6.2), it was observed that certain distinctive differences among subjects within the CLBP group could be established. This triggered motivation to inspect all profiles in more detail with the intention to consequently provide patients subgrouping and clustering where possible. Profiles for all twenty-nine CLBP patients were constructed and presented in Fig. 6.3.

One of the initial challenges was whether we can provide such a procedure or "tool", intuitive enough for medical experts, that can easily relate (even visually) the modeling results with the actual individual patient's LBP status. For this reason, as the first step, a visual (manual) approach was applied to remove the formal mathematical approach sometimes masking or aggravating human-readable and intuitive inference. Later in the text, more formal mathematical techniques and procedures will be applied to validate the respective outcomes.

To support this idea, a high-level (rough) subgrouping based on visual inspection and geometric contour shape visual matching was performed with the aim to isolate those most distinctive subgroups, if and where possible. This attempt resulted in five proposed subgroups that were believed to be most distinctive by visual inspection criteria, altogether grouping 15 out of 29 CLBP subjects, with additionally annotating 4 CLBP profiles as candidates to be assigned to one of these 5 isolated subgroups (Fig. 6.3). Thus, the majority of CLBP patients, at hand, were assigned to one of the subgroups, whereas for the remaining this approach was not providing decisive insights. The profiles relating to each of the respective subgroups were annotated with surrounding dashed boxes colored in one of the following five colors: green, red, blue, purple, and orange. The candidate profile boxes were colored with the light version of the respective colors, namely, light green, light red, etc.

In the following subsections, the proposed subgroups are analyzed in more detail.

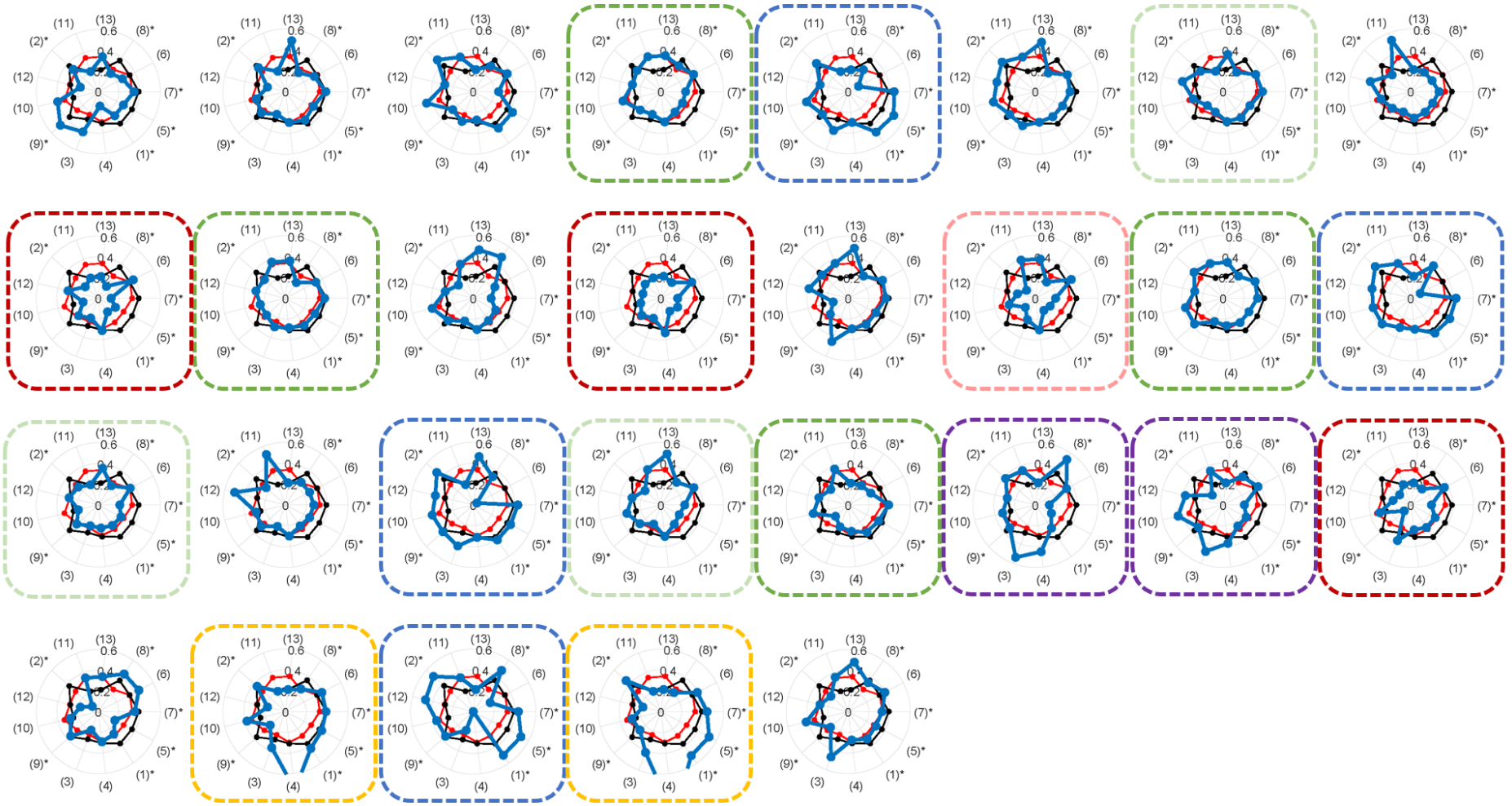


Figure 6.3: CLBP profiles for all twenty-nine patients are shown (blue). Subgroups are annotated with respective colored dashed boxes in green, red, blue, purple, and orange for *Regular CLBP pattern*, *Inhibited pattern*, *Type A coordination imbalance pattern*, *Type B coordination imbalance pattern*, and *Exceeding coordination pattern*, respectively. Profiles are stacked row-wise, meaning, the first profile in each row corresponds to the 1st, 9th, 17th, and 25th CLBP patient. Candidate profiles are annotated with light versions of respective colors. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively.

6.3.1 Regular CLBP pattern

As the first subgroup, those profiles most closely following the median CLBP profile were isolated, subgrouped and named *Regular CLBP pattern* patients (Fig. 6.4). In Fig. 6.3 these profiles were annotated with greendashed boxes and corresponded to subjects: 4, 10, 15, and 21. The candidate profiles bounded with light green dashed boxes are: 7, 17, and 20.

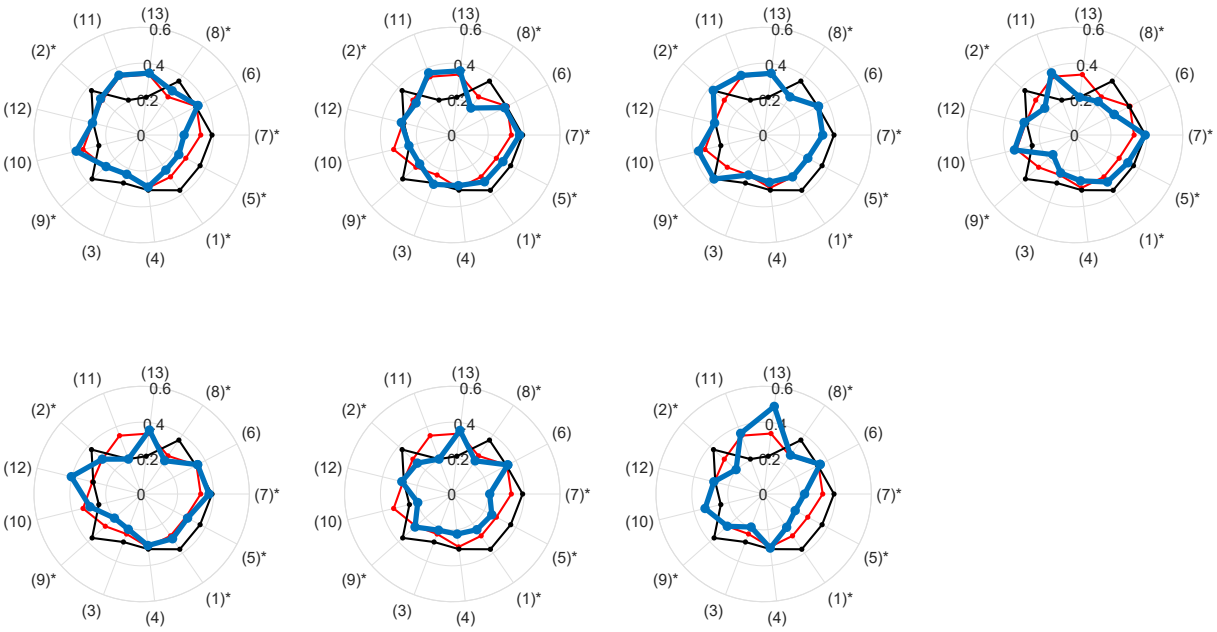


Figure 6.4: CLBP profiles for *Regular pattern* are shown (blue) for subjects 4, 10, 15, and 21 (top) and candidate subjects 7, 17, and 20 (bottom), from left to right, respectively. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.2. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

Further analysis will reveal whether these profiles can be correlated to some of the well-established results in the realm of LBP studies and beliefs about how neuromotor control and myoelectric activity are expressed for CLBP patients.

6.3.2 Inhibited pattern

For the next CLBP profile type, a set of quite distinguished patients profiles was selected and named *Inhibited pattern* as shown in Fig. 6.5. These profiles are strongly characterized by reduced feature component values both compared to median CLBP and HS profiles, thus named "inhibited". In Fig. 6.3 these profiles were annotated with reddashed boxes and corresponded to subjects: 9, 12, and 24. The candidate profile bounded with light red dashed boxes is 14.

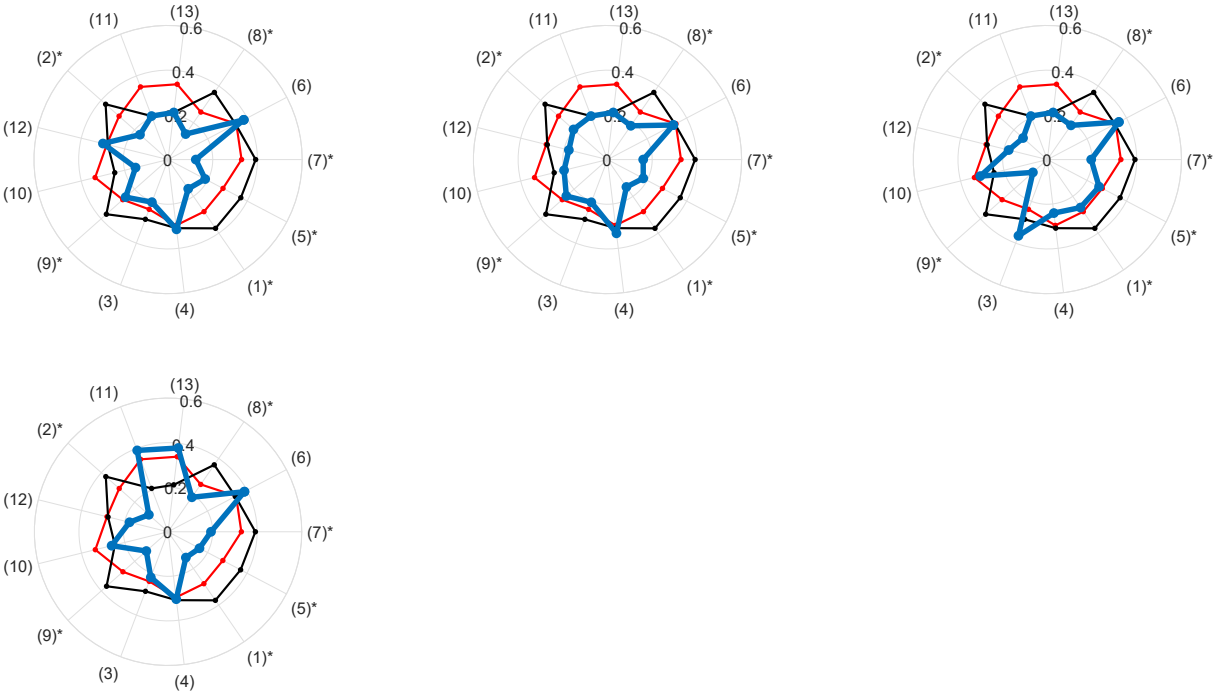


Figure 6.5: CLBP profiles for *Inhibited pattern* are shown (blue) for subjects 9, 12, and 24 (top) and candidate subject 14 (bottom), from left to right, respectively. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.2. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

6.3.3 Type A coordination imbalance pattern

As the third CLBP profile type, a set of also well distinguished patients profiles was selected and named *Type A coordination imbalance pattern* as shown in Fig. 6.6. For this profile, significant coordination features relating to more complex myoelectric activity in the LLES region were noticed, much higher than in the other CLBP patients, moreover exceeding standard complex coordination activity in HS subjects as well. On the other hand, at the same time, an extreme reduction in coordinated muscle force is present. In Fig. 6.3 these profiles were annotated with bluedashed boxes and corresponded to subjects: 5, 16, 19, and 27. For this profile type, no additional candidate profiles were selected.

6.3.4 Type B coordination imbalance pattern

As the forth CLBP profile type, a sort of counterpart for *Type A coordination imbalance* part was detected and named *Type B coordination imbalance pattern* as shown in Fig. 6.7. This pattern is characterized by almost opposite coordination patterns in the LLES region. Namely, more complex coordination activity features are reduced where simple coordination of the level of the exerted force (energy) between the left and right side is put in balance. In Fig. 6.3 these profiles were annotated with purpledashed boxes and corresponded to subjects: 22 and 23.

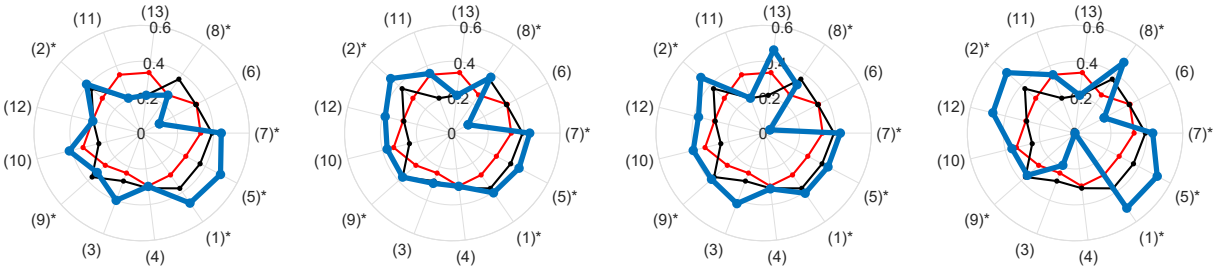


Figure 6.6: CLBP profiles for *Type A coordination imbalance pattern* are shown (blue) for subjects 5, 16, 19, and 27, from left to right, respectively. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.2. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

For this profile type, no additional candidate profiles were selected.

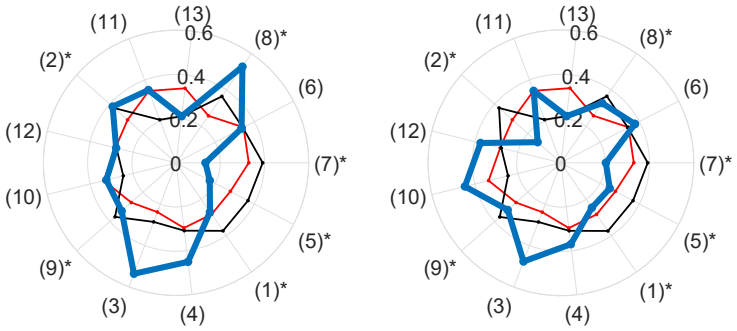


Figure 6.7: CLBP profiles for *Type B coordination imbalance pattern* are shown (blue) for subjects 22 and 23, from left to right, respectively. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.2. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

6.3.5 Exceeding coordination pattern

As the fifth CLBP profile type, a well-distinguished profile relating to non-standard exceeding muscle activity in the LLES region was detected and named *Exceeding coordination pattern* as shown in Fig. 6.8. This pattern is characterized by most of the feature metrics in the LLES region exceeding not only standard (median) CLBP values, but also exceeding, or at worst matching, the respective metrics for the HS group. In Fig. 6.3 these profiles were annotated with orangedashed boxes and corresponded to subjects: 26 and 28. For this profile type, no additional candidate profiles were selected.

6.4 Clustering

As a further step toward CLBP patients subgrouping, formal mathematical procedures and machine learning techniques were employed. Such an approach was required to establish objective

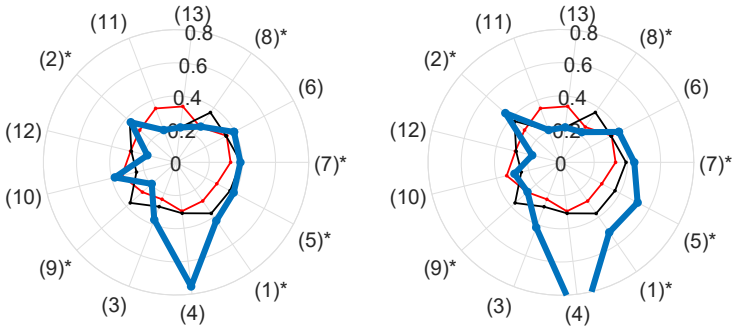


Figure 6.8: CLBP profiles for *Exceeding coordination pattern* are shown (blue) for subjects 26 and 28, from left to right, respectively. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.2. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

and quantifiable measures and metrics, alongside cross-checking the earlier proposed subgroups and validating whether meaningful correlation and connection between visually selected profiles are formal mathematical clustering outcomes can be established.

In the following subsections, two different unsupervised clustering techniques were employed and examined: self-organizing maps (SOM) and hierarchical clustering.

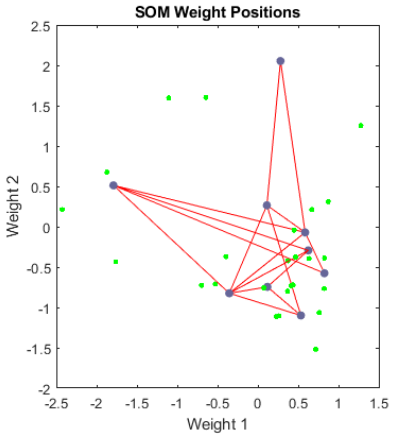
6.4.1 Self-organizing maps (SOM)

SOM is a neural clustering method used for pattern discovery, often deemed to exhibit immense discovery power and as such used in many scenarios for exploratory data analysis (EDA) [170] or time series mining with applications in the medical domain as well [171]. Moreover, applied by Riveros *et al.* [172] to categorize data for diagnosing spinal pathology and predict LBP.

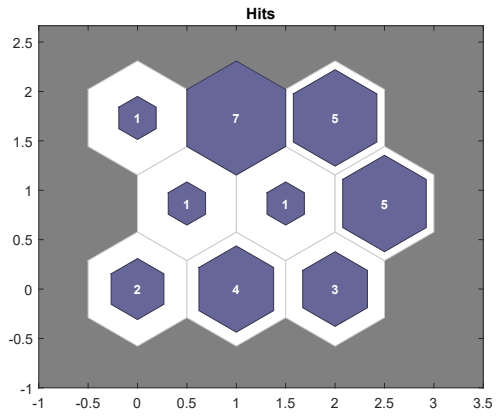
SOM can be used to classify information and reduce the variable number of complex problems. It also allows the visualization of information via a two-dimensional mapping. A SOM is a simple layer neural network with N neurons distributed in space in a grid mode. The characteristic of the SOM technique is to produce a lower-dimensional representation of data alongside preserving the topological structure of data, where these representations are in most applications rectangular, although hexagonal representations are also common. The algorithm learning process is based on the similarity between the input vector and the distance between neurons. For similarity measures, Euclidean methods are used, and it is also necessary to use correlation, directional cosine, and block distances [172].

In this PhD thesis, the Neural Clustering tool as part of Matlab software package [144] was used to perform SOM clustering with out-of-the-box parameters. Topology was in hexagonal representation with a 3x3 two-dimensional map. This number of outputs (nine) was chosen based on preliminary visual subgrouping (section 6.3) where it was reasonable to assume that number of potential subgroups to detect will be higher than five, but less than nine.

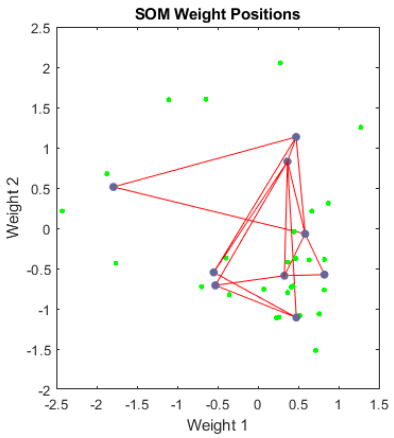
One of the challenges pertained to SOM clustering is that the resulting topology grouping (i.e hits) largely depends on different initial conditions and sampling, where each (re)train provided different results, as can be observed from Fig. 6.9b and Fig. 6.10b. Similarly, the position of nodes (implying SOM model weights) also differs across different iterations as seen in Fig. 6.9a and Fig. 6.10a. Therefore, this technique was not straightforward to establish consistent insight into the clustering results and subgrouping, moreover as there was not a direct linkage between the groupings hits and what actual subjects that are part of that grouping (i.e., it was not explicitly known which subject corresponds to which subgroup).



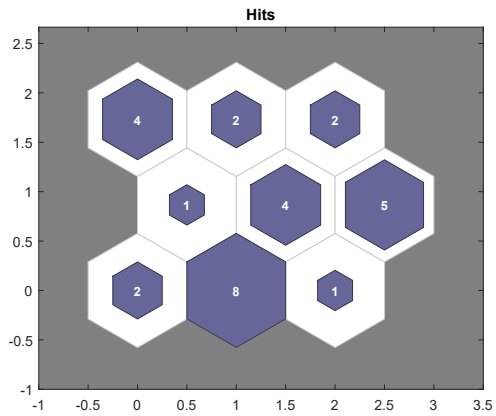
(a) Train A: SOM weight positions.



(b) Train A: SOM hits in 3x3 hexagonal topology.



(a) Train B: SOM weight positions.



(b) Train B: SOM hits in 3x3 hexagonal topology.

6.4.2 Hierarchical clustering

Another approach to CLBP subject's subgrouping was via hierarchical clustering (sometimes also called hierarchical cluster analysis - HCA). Hierarchical clustering, as an unsupervised clustering technique, groups similar objects (samples) into clusters, where it either uses an agglomerative ("bottom-up") or divisive ("top-down") strategy to build the hierarchy (dendrogram). The bottom of the hierarchy is a set of clusters that differ from each other, where clusters

closer in the hierarchy tend to contain more similar objects (based on some similarity metric criteria) [173]. One of the advantages of this approach is that it does not require the number of clusters to be defined in front (as an input into the algorithm), unlike for k-means clustering method. Moreover, the dendrogram enables an easy visual insight into the number of clusters and their relations, thus tends to be more informative than flat clusters as ones returned by k-means.

In this PhD thesis, a simple setup for the respective hierarchical clustering was employed. An agglomerative clustering, based on default Euclidean distance, and commonly used Ward's linkage method [174], based on inner squared distance (minimum variance algorithm) which is appropriate for Euclidean distances only, were utilized [144]. There are no clear theoretical justifications for the choice of linkage criteria, but Ward's method is considered the sensible default. Ward's method finds which objects to group based on reducing the sum of squared distances of each object against the mean object in a cluster. This concept of distance is in many cases deemed appropriate as it matches the standard assumptions of how differences between groups in statistics are calculated (e.g., ANOVA, MANOVA). For the clustering parameters, no predefined number of clusters were set, but *inconsistency coefficient* criterion with value 3 was selected and *depth* parameter, for computing inconsistent values by looking to a depth D below each node, was set to value 5 [144]. The inconsistency coefficient is calculated for each link by comparing its height with the average height of other links at the same level of the hierarchy. The larger the coefficient, the greater the difference between the objects connected by the link [175].

For the given clustering setup, a respective dendrogram has been created and is shown in Fig. 6.11. The respective endpoint and bottom level clusters are compared with the preliminarily defined clusters isolated by visual inspection (as described in section 6.3) and annotated with the respective cluster color coding. Insightfully, the hierarchical clustering has provided almost identical subgrouping for the respective patients compared to the initial manual approach which was based on only employing the visual shape matching in the radial plot representation. Moreover, even the candidate profiles fitted into the same clusters. The only exception was observed for subject 17 where the candidate profile was set to be the *Regular CLBP pattern*, whereas the hierarchical clustering has grouped it in the same cluster hierarchy as the subject described by the *Inhibited pattern* profiles.

Overall, the results and insights gathered here were now considered the strong supporting point that implied that indeed an intuitive visual inspection approach can be related to more formal mathematical background and well-known grouping (clustering) techniques. This proved further that the proposed feature modeling and extracted model components, describing differences between HS and CLBP groups, are suitable for deeper inference into the CLBP group.

An additional noteworthy aspect was that, unlike the SOM clustering procedure, hierarchical

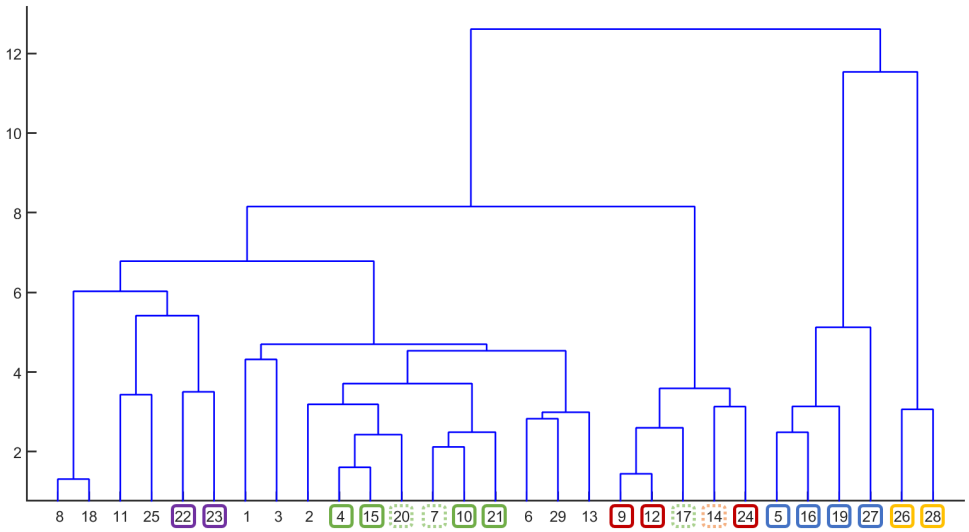


Figure 6.11: Hierarchical clustering dendrogram for CLBP profiles based on thirteen NCA components, with preliminary subgroup annotations and color coding for five subgroups selected, as described in section 6.3.

clustering was consistently providing the same grouping results (on this data set) irrespective of repeated (re)runs of the respective algorithm. Potential challenges when applying hierarchical clustering are mostly related to time-complexity for large data sets, but this was not considered a significant disadvantage for this particular use case due to relatively small data set sizes that are realistically expected in clinical applications, with a max couple of hundred subjects (samples) available at hand.

6.5 Subgroups analysis and validation

The insightful results obtained in section 6.4.2 were further exploited to consolidate the CLBP patients' subgrouping and to provide some quantitative validation of the given results.

Firstly, the candidate profiles were assigned to the proper subgroup profiles based on the hierarchical insights. Both candidate profiles 14 and 17 were assigned to *Inhibited pattern* subgroup. Further, it was opted to expand our subgroups with two additional groups, based on visual inspection and linkage distances among objects in the cluster. The consolidated profiles and new clusters are given in Fig. 6.15 and Fig. 6.14.

For new additional two groups, one consists of subjects 6, 13, and 29 (blue-graycolor coded) and is named *Near-regular pattern* (Fig. 6.12), and the other group consists of subjects 8 and 18 (light-graycolor coded) and is named *Co-activation imbalance pattern* (Fig. 6.13).

With this step, 25 out of 29 CLBP subjects were assigned to one of seven subgroups (clus-

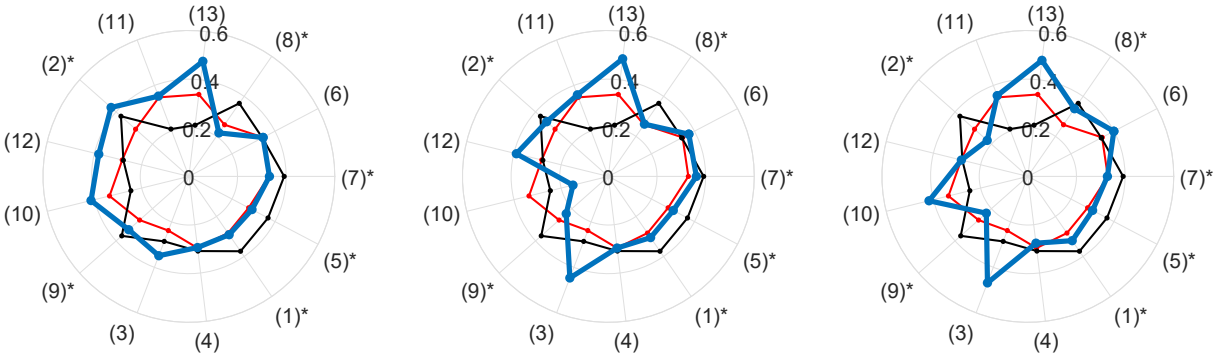


Figure 6.12: CLBP profiles for *Near-regular pattern* are shown (blue) for subjects 6, 13 and 29, from left to right, respectively. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.2. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

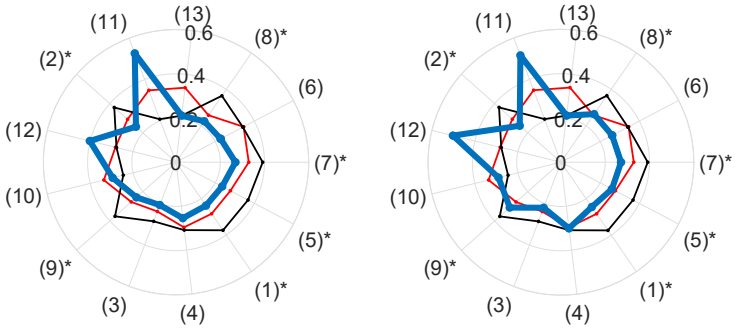


Figure 6.13: CLBP profiles for *Co-activation imbalance pattern* are shown (blue) for subjects 8 and 18, from left to right, respectively. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively. NCA components are labelled with (1) to (13) and referencing to Table 6.2. Components with * indicate the rejected Wilcoxon rank sum test null hypothesis at 5% significance level.

ters). The remaining 4 subjects (profiles) were considered non-standard and in this step were left outside, but with the potential to assign them more elaborately to one of the existing or new subgroups in the subsequent steps. It can be also noticed that the profiles relating to *Regular CLBP pattern* are, individually as a subgroup, the most frequent ones. On the intuitive level, this makes sense.

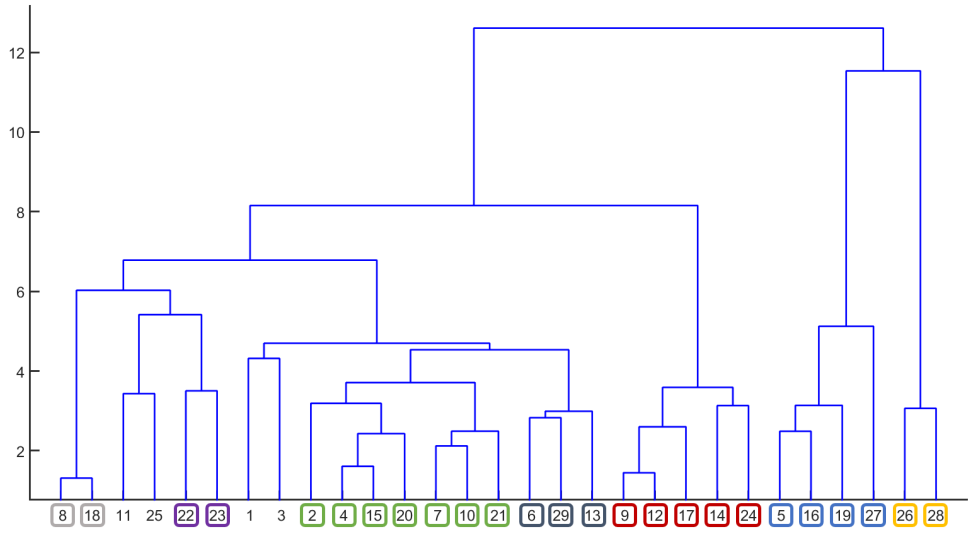


Figure 6.14: Hierarchical clustering dendrogram for CLBP profiles based on thirteen NCA components. Endpoints are color coded as per the consolidated subgrouping (Fig. 6.15) with seven distinct subgroups. CLBP subjects 1, 3, 11, and 25 were not assigned to any of the given subgroups.

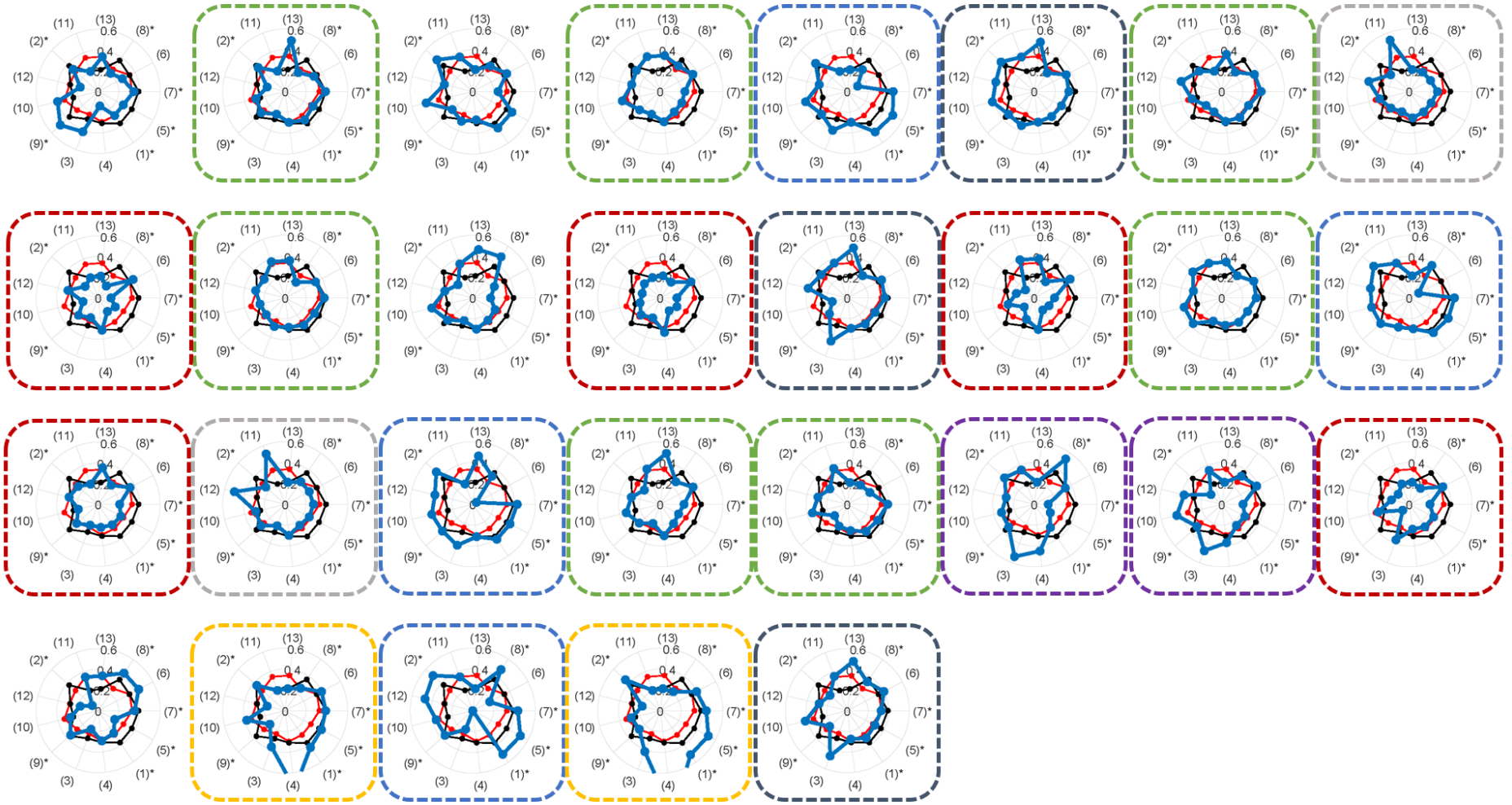


Figure 6.15: CLBP profiles for twenty-nine patients are shown (blue) alongside seven consolidated subgroups. Subgroups are annotated with respective colored dashed boxes in green, red, blue, purple, orange, blue-gray, and light-gray for *Regular CLBP pattern*, *Inhibited pattern*, *Type A coordination imbalance pattern*, *Type B coordination imbalance pattern*, *Exceeding coordination pattern*, *Near-regular pattern*, and *Co-activation imbalance*, respectively. Profiles are stacked row-wise, meaning, the first profile in each row corresponds to the 1st, 9th, 17th, and 25th CLBP patient. CLBP profiles 1, 3, 11, and 25 were not assigned to any of the seven subgroups. Median component values across all HS and CLBP subjects are shown in black and red solid contours, respectively.

6.5.1 Relationship with metadata

In this subsection, the analysis of potential relationships between subgroup types and respective metadata was performed. Following metadata were included: BMI, exercise duration (time-to-fatigue), ODQ, RDQ, VAS, and FFK. FFK was measured for both the left (FFK sin) and right (FFK dex) side, where across all CLBP subjects, as well as for HS subjects, no significant differences between left and right-side values were observed, whereas for the RLBP group such difference is observable (Table 3.1). To supplement this analysis, the average values for the respective metadata were also provided overall for CLBP and HS groups, separately, calculated across all subjects (meaning, in this case, both groups were treated as homogeneous). All these results are presented in Table 6.3.

As a baseline observation, it can be noticed that the HS group exhibits the lowest values for LBP and pain level detection questionnaires (ODQ, RDQ, VAS), demonstrated long exercise durations ($\sim$140 s on average) and high FFK value ($\sim$80°). These values can be considered as expected for a healthy subject's cohort. Looking deeper into CLBP subgroups, some patterns might be recognized, although with caution, as even the largest subgroup sample (namely, *Regular CLBP pattern*) contains only seven subjects, thus statistical significance cannot be assumed. Nevertheless, some of the observations were:

- (I) *Regular CLBP pattern* demonstrates below average exercise duration and above average ODQ compared to both overall CLBP and HS mean values.
- (II) *Inhibited pattern* subjects exhibited the shortest exercise durations, noticeably differing from other subgroups. Also, the average FFK values were in the lower end among all subgroups.
- (III) *Type A coordination imbalance pattern* subjects are characterized by long exercise durations ($\sim$150 s on average), high FFK values, and low ODQ, RDQ, VAS values, thus, much resembling the HS group when considering these metadata only.
- (IV) *Type B coordination imbalance pattern* is also characterized by long exercise durations (>150 s on average), however with somewhat lower FFK values (65°-70°) and the highest ODQ, RDQ, and VAS values.
- (V) *Exceeding coordination pattern* is primarily characterized by relatively high FFK values ($\sim$80°) and very low RDQ (0-1). However, it needs to be noted that only two subjects were part of this subgroup.
- (VI) *Near-regular pattern* is characterized by longer exercise durations ($\sim$120 s on average), higher FFK values ($\sim$80°) and similar ODQ and RDQ values, compared to *Regular CLBP pattern*.
- (VII) *Co-activation imbalance pattern* is characterized by short exercise durations ($\sim$70 s on average), the lowest FFK values ($\sim$60°), and among the highest ODQ, RDQ, and VAS values.

- *Non-specified* section of unassigned profiles has some common characteristics. Namely, very long exercise durations, hitting the exercise duration limit of 180 s alongside with majority of these profiles having above average FFK values, comparable with HS group, where 2 of 4 subjects have reported low ODQ, RDQ, VAS values. For these 2 subjects, the respective metadata might be easily misinterpreted as to be corresponding to the HS group.

Overall results and insights suggest that there is a certain connection between the proposed feature modeling subgrouping (as proposed in this PhD thesis) and the analyzed metadata. Namely, when metadata were grouped according to the subgroup profile patterns, some of the mean values for the respective measures showed up as distinctive values specific for that very subgroup (Table 6.3). One of the interesting insights revealed with this analysis is for the set of *non-specified* subjects. Here, above-average values for exercise duration and FFK are observed, pointing to the certain similarity that is closer to the HS group than to the CLBP group itself. Additional insight into the potential relationships was provided by the HCA dendrogram calculated specifically for the normalized metadata parameters set as shown in Fig. 6.16. It can be noticed that HCA metadata grouping very weakly coincides with the subgrouping based on CLBP profiles, where such loose matching is sporadically observed only for *Regular CLBP pattern* and *Inhibited pattern* subgroups. For other subgroups, no indication for subgrouping as such could be observed (small subgroup sizes also impact this type of visual dendrogram inspection).

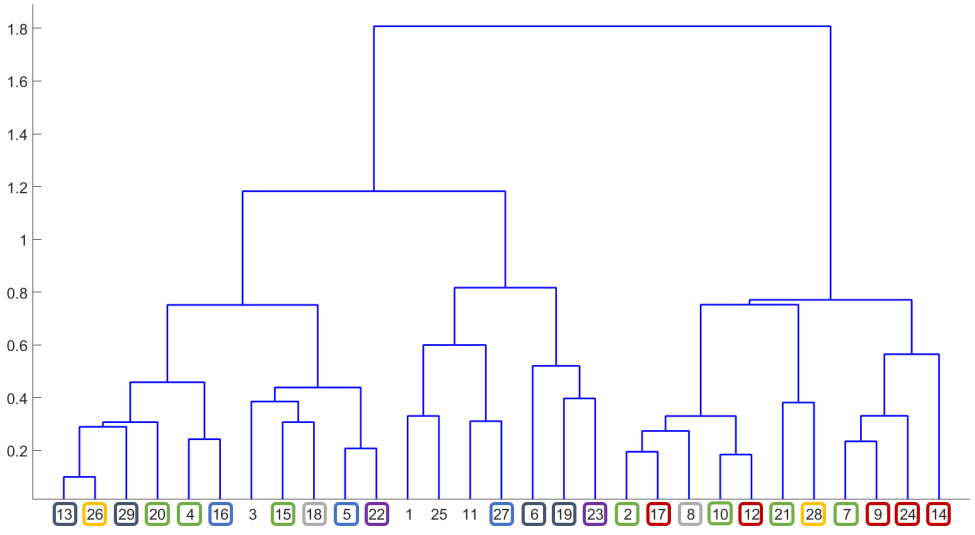


Figure 6.16: Hierarchical clustering dendrogram for CLBP group metadata. Endpoints are color coded as per the consolidated subgrouping (Fig. 6.15) with seven distinct subgroups. CLBP subjects 1, 3, 11, and 25 were not assigned to any of the given subgroups.

Table 6.3: Metadata for all CLBP subjects grouped per subgroup types, with respective mean values for each subgroup (except for Non-specified ones), and mean values for all subjects across CLBP group and HS group, separately. Exercise duration, BMI, FFK dex (right), FFK sin (left), ODQ, RDQ, and VAS measures are shown. Subgroup patterns are color coded as per the consolidated subgrouping (Fig. 6.15) with seven distinct subgroups.

Subgroup	Subject ID	Exercise duration (s)	BMI	FFK dex (°)	FFK sin (°)	ODQ	RDQ	VAS
Regular CLBP pattern (I)	2	88.00	27.47	68.00	68.00	30.00	8.00	4.00
	4	111.00	20.90	74.00	70.00	4.00	4.00	0.00
	7	63.00	26.85	76.00	76.00	6.00	5.00	3.00
	10	60.00	28.46	65.00	64.00	24.00	10.00	5.00
	15	112.00	24.58	60.00	60.00	18.00	4.00	2.00
	20	85.00	23.65	77.00	79.00	34.00	1.00	2.00
	21	56.00	24.40	85.00	85.00	24.00	8.00	6.00
	Mean	82.14	25.19	72.14	71.71	20.00	5.71	3.14
Inhibited pattern (II)	9	32.00	27.78	71.00	71.00	14.00	4.00	2.00
	12	49.00	29.05	55.00	61.00	32.00	10.00	4.00
	14	27.00	27.77	87.00	87.00	18.00	12.00	3.00
	17	75.00	26.32	65.00	70.00	16.00	8.00	3.00
	24	43.00	28.41	69.00	71.00	4.00	2.00	0.00
	Mean	45.20	27.87	69.40	72.00	16.80	7.20	2.40
Type A coordination imbalance (III)	5	125.00	23.82	79.00	79.00	16.00	6.00	2.00
	16	120.00	23.15	81.00	81.00	0.00	0.00	0.00
	19	180.00	24.11	71.00	71.00	2.00	7.00	1.00
	27	180.00	23.18	90.00	90.00	12.00	6.00	4.00
	Mean	151.25	23.57	80.25	80.25	7.50	4.75	1.75
Type B coordination imbalance (IV)	22	140.00	23.59	80.00	70.00	18.00	3.00	3.00
	23	177.00	24.34	60.00	60.00	20.00	14.00	0.00
	Mean	158.50	23.96	70.00	65.00	19.00	8.50	1.50
Exceeding coordination pattern (V)	26	95.00	27.06	78.00	80.00	18.00	0.00	1.00
	28	86.00	27.10	85.00	79.00	10.00	1.00	6.00
	Mean	90.50	27.08	81.50	79.50	14.00	0.50	3.50
Near-regular pattern (VI)	6	128.50	25.98	85.00	87.00	22.00	12.00	1.00
	13	94.00	27.73	77.00	77.00	10.00	1.00	1.00
	29	120.00	24.41	82.00	82.00	26.00	2.00	0.00
	Mean	114.17	26.04	81.33	82.00	19.33	5.00	0.67
Co-activation imbalance (VII)	8	64.00	25.93	55.00	55.00	26.00	7.00	3.00
	18	80.00	25.01	67.00	67.00	12.00	2.00	4.00
	Mean	72.00	25.47	61.00	61.00	19.00	4.50	3.50
Non-specified	1	180.00	26.02	103.00	103.00	0.00	4.00	0.00
	3	130.00	25.64	60.00	64.00	18.00	9.00	4.00
	11	180.00	28.01	70.00	80.00	14.00	4.00	4.00
	25	179.00	24.69	85.00	85.00	4.00	1.00	1.00
CLBP subjects (all)	**Mean**	105,05	25.70	74.48	74.90	15.59	5.34	2.38
Healthy subjects (all)	**Mean**	139.95	25.35	80.43	80.59	7.14	1.86	1.65

6.5.2 Quantitative analysis and proposed metrics

In this section, so far collected insights and inputs are set to be quantified more concretely. Namely, one of the tasks at hand could be - if having a new CLBP subject with a respective profile (i.e. respective set of NCA components values), what would be a proper procedure or metric to validate which cluster or subgroup new sample is most likely to be assigned to? Visual inspection and manual subgrouping can provide satisfactory results to some extent, but such an approach is obviously too subjective, time-consuming, and prone to errors. An automated subgrouping (clustering) method is expected to be provided. In that course, hierarchical clustering is a more formal mathematical approach, with a very convenient dendrogram visualization of the resulting clustering result. However, we are still missing some sort of metric that can quantify in an easy way how similar an individual's profile is to any of the subgroup patterns.

Here, an approach to tackle this task is proposed. As the first step, for each subgroup, an averaged pattern profile was calculated. This way, seven reference templates were constructed. Though, we need to have in mind that a small number of subjects in each subgroup poses an additional challenge in terms of statistical significance and relevance of such calculated templates. Meaning, the risk of dealing with samples treated as "outliers" is thus somewhat inherent. The resulting subgroup template patterns are shown in Fig. 6.17.

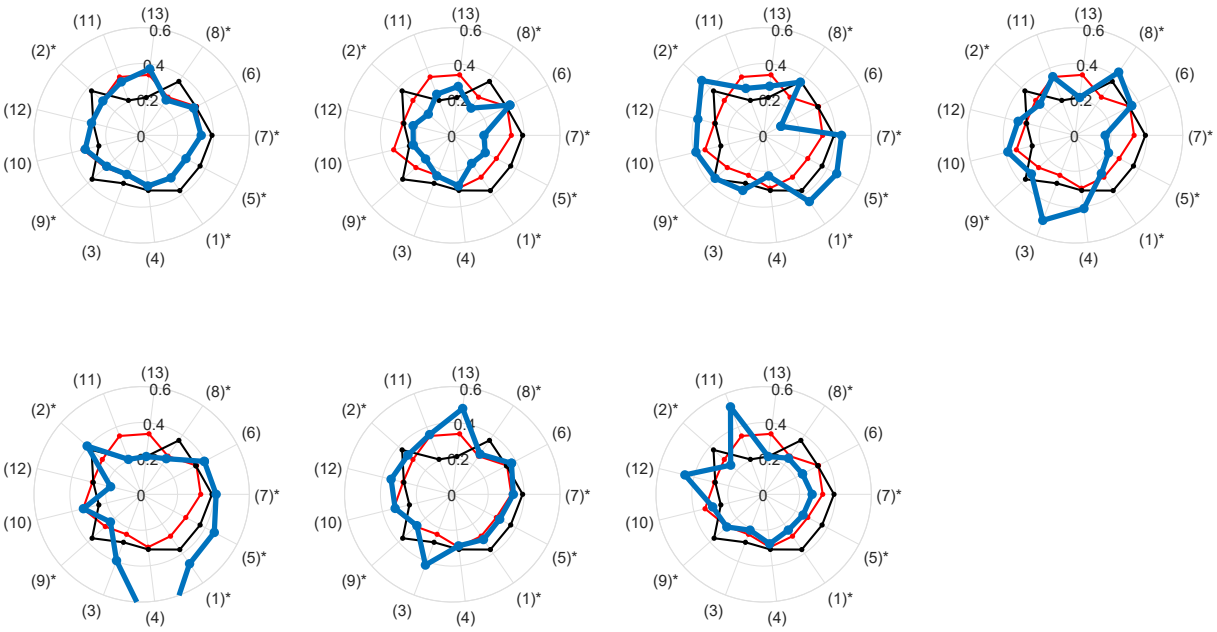


Figure 6.17: Averaged profile templates for respective subgroup patterns (I to VII). Top row, four templates for patterns I to IV, namely, for *Regular CLBP pattern*, *Inhibited pattern*, *Type A coordination imbalance pattern*, and *Type B coordination imbalance pattern*, from left to right, respectively, are shown. Bottom row, three templates for patterns V to VII, namely, for *Exceeding coordination pattern*, *Near-regular pattern*, and *Co-activation imbalance pattern*, from left to right, respectively, are shown.

In the second step, two different distance metrics are defined. Metric A, as the sum of

differences between an averaged template and the respective subject's profile across all profile components c (6.1). Metric B is similarly defined where the sum of squared differences between an averaged subgroup's template and the respective subject's profile is calculated (6.2).

$$M_{A_{ij}} = \sum_c \bar{P}_{i,c} - s_{j,c} \tag{6.1}$$

$$M_{B_{ij}} = \sum_c \left(\bar{P}_{i,c} - s_{j,c}\right)^2 \tag{6.2}$$

where c stands for NCA components, $\bar{P}_i$ stands for an averaged template corresponding to pattern P_i, i represents one of the seven isolated patterns, and s_j denotes the j-th subject at test.

Now, once metrics are defined, the existing data set and respective subgroups (patterns) are verified in terms of calculating distances between each subject and each subgroup template, for both metrics. Results are presented in Table 6.4. The obtained results are very satisfactory, where Metric A (6.1) confirmed 18 out of 25 subgroup assignments when compared to clustering results that are suggested by the HCA dendrogram. For Metric B (6.2), these results were even better, all 25 subjects were assigned to the same subgroups as initially suggested by hierarchical clustering. Interestingly, the *non-specified* subjects are shown to be closest to the *Regular CLBP pattern* (I) and *Near-regular pattern*, as per Metric A and Metric B calculations. However, it is deemed not plausible to forcibly include this subset into the *Regular CLBP pattern* subgroup. Namely, average Metric B distances for the closest subgroup are <0.05, whereas for this subgroup these distances were 3-4 times larger, thus leaving this group out, especially having in mind the corresponding metadata as well (e.g., very high endurance times), together with very high FFK values (for 3 out of 4 subjects).

These results are deemed important as we see a "full circle" being made, i.e., different angles and approaches provided consistent results and insights - namely, the subgroups introduced, by visually inspected CLBP models, were confirmed and additionally expanded by out-of-the-box hierarchical clustering machine learning technique. Subsequently, such clustering results were examined by introducing a new *construct*, i.e., averaged subgroup templates, derived from the proposed CLBP pattern models, and validated by defining a simple distance metric based on the sum of squared differences. The results obtained with this last step suggested that the selected subgroups indeed possess certain characteristics that make them distinctive against each other and relatively easy to label by employing simple metrics. These distinctive characteristics were to some extent observed even in the part of the analysis that involved metadata inputs. Moreover, it is important to notice that the subgrouping techniques for HCA and here proposed template-based matching differ in nature. Hierarchical clustering builds clusters bottom-up in an agglomerative manner with no predefined references or templates defined in front, where individual observations are treated as separate clusters (endpoints), and then in an iterative manner

two most similar endpoints (clusters) are merged into new (larger) clusters based on some similarity metric. This iterative procedure is then repeated until all clusters are merged into one big cluster. On the other hand, the approach proposed in this PhD thesis has established in front a set of knowledge-based predefined templates (expanded by two additional subgroups based on dendrogram insights) and each observation is assigned to a respective subgroup based on the smallest sum of squared distances between the respective observation (subject) and the averaged pattern template. Here presented approach provided certain flexibility in defining which profile corresponds to which subgroup (pattern). Namely, it could have been observed in Fig. 6.14 that some observations, as for *Regular CLBP pattern* (green) and *Inhibited pattern* where extending across several hierarchy levels, whereas some observations, like those corresponding to *Type B coordination imbalance pattern*, *Exceeding coordination pattern*, or *Co-activation imbalance* where kept within single hierarchy level. The presented pattern-wise grouping is thus considered as a new *construct* introduced by the initial visually based inspection of CLBP profiles at hand. Nonetheless, it is fair noticing that in the background same/similar concepts of quantifying similarity between two entities were employed, namely both Ward's linkage for HCA and Metric B (6.2) for template-matching are based on minimizing the sum of squared distances from the averaged cluster (pattern) representative.

As a continuation of these considerations, new measure defined by (6.2) was introduced as a new plausible *CLBP Pattern Distance* (CPD) metric for assigning any new CLBP subject, j, (modelled with the features presented in this PhD thesis) to any of the seven CLBP subgroups, P_i, $i = 1..7$ (section 6.5), based on minimum CPD distance as annotated in (6.3).

$$\min_i CPD_i\left(s_j\right) \rightarrow s_j \in P_i \tag{6.3}$$

The presented outcomes are strongly suggesting that the more homogeneous subgroups within the CLBP population exist and that the subgrouping can be achieved exploiting the proposed LBP features modeling. The next key aspect is to provide meaningful medically based interpretation for each of the proposed subgroups. These aspects will be tackled further in the text.

Table 6.4: Quantification of distance (similarity) measures among each subject (ID) and the respective pattern's (I to VII) mean profile calculated by two metrics (A and B) are shown. Pattern annotations (I to VII) are defined as in section 6.5.1. Hit columns show the closest (most similar) pattern in terms of metrics A and B, for each subject respectively. Hits colored in green denote a successful match between hierarchical clustering results and respective labeling based on metrics A and B, given by equations (6.1) and (6.2). Subjects not assigned to any subgroup (pattern) are initially annotated with pattern label 0.

Pattern	ID	Distance to pattern (Metric A)								Distance to pattern (Metric B)							
		I	II	III	IV	V	VI	VII	Hit	I	II	III	IV	V	VI	VII	Hit
I	2	-0.07	-0.91	0.61	0.29	0.92	0.50	-0.17	I	0.05	0.17	0.23	0.33	0.50	0.10	0.29	I
I	4	0.02	-0.83	0.69	0.38	1.00	0.59	-0.08	I	0.02	0.08	0.26	0.14	0.57	0.08	0.10	I
I	7	0.03	-0.82	0.70	0.39	1.01	0.60	-0.07	I	0.03	0.12	0.22	0.27	0.53	0.09	0.16	I
I	10	-0.05	-0.89	0.63	0.31	0.94	0.53	-0.15	I	0.02	0.11	0.22	0.23	0.50	0.06	0.11	I
I	15	-0.26	-1.11	0.41	0.10	0.72	0.31	-0.36	IV	0.02	0.15	0.17	0.19	0.55	0.07	0.12	I
I	20	0.11	-0.73	0.79	0.48	1.10	0.69	0.02	VII	0.05	0.10	0.39	0.23	0.69	0.10	0.15	I
I	21	0.22	-0.62	0.90	0.59	1.21	0.80	0.13	VII	0.06	0.14	0.20	0.25	0.55	0.16	0.09	I
II	9	1.03	0.19	1.71	1.40	2.02	1.61	0.94	II	0.17	0.03	0.57	0.29	0.72	0.30	0.21	II
II	12	1.04	0.20	1.72	1.41	2.03	1.62	0.95	II	0.14	0.02	0.51	0.26	0.64	0.29	0.22	II
II	14	0.79	-0.05	1.46	1.15	1.78	1.36	0.69	II	0.13	0.05	0.62	0.30	0.77	0.23	0.21	II
II	17	0.53	-0.31	1.21	0.90	1.52	1.11	0.44	II	0.05	0.04	0.33	0.24	0.67	0.14	0.17	II
II	24	0.82	-0.02	1.50	1.19	1.81	1.40	0.73	II	0.14	0.05	0.46	0.23	0.63	0.23	0.27	II
III	5	-0.50	-1.35	0.17	-0.14	0.48	0.07	-0.60	VI	0.24	0.44	0.05	0.36	0.44	0.26	0.37	III
III	16	-0.65	-1.50	0.02	-0.29	0.33	-0.08	-0.75	III	0.20	0.44	0.02	0.30	0.54	0.24	0.22	III
III	19	-0.81	-1.66	-0.14	-0.45	0.17	-0.24	-0.91	III	0.25	0.52	0.07	0.39	0.59	0.21	0.41	III
III	27	-0.73	-1.58	-0.06	-0.37	0.25	-0.16	-0.83	III	0.39	0.69	0.12	0.61	1.05	0.44	0.42	III
IV	22	-0.47	-1.32	0.20	-0.11	0.51	0.10	-0.57	VI	0.30	0.33	0.41	0.04	0.53	0.25	0.33	IV
IV	23	-0.26	-1.10	0.42	0.11	0.73	0.32	-0.35	IV	0.17	0.19	0.37	0.04	0.56	0.16	0.18	IV
V	26	-0.69	-1.54	-0.02	-0.33	0.29	-0.12	-0.79	III	0.34	0.45	0.46	0.34	0.03	0.38	0.54	V
V	28	-1.28	-2.12	-0.60	-0.91	-0.29	-0.70	-1.37	V	0.75	0.91	0.78	0.72	0.03	0.75	0.98	V
VI	6	-0.67	-1.51	0.00	-0.31	0.32	-0.10	-0.77	III	0.07	0.25	0.19	0.22	0.56	0.03	0.19	VI
VI	13	-0.47	-1.31	0.21	-0.11	0.52	0.11	-0.57	IV	0.11	0.22	0.29	0.24	0.58	0.04	0.22	VI
VI	29	-0.58	-1.43	0.09	-0.22	0.40	-0.01	-0.68	VI	0.10	0.22	0.29	0.15	0.57	0.03	0.26	VI
VII	8	0.26	-0.58	0.94	0.62	1.25	0.83	0.16	VII	0.11	0.16	0.31	0.24	0.75	0.20	0.01	VII
VII	18	-0.07	-0.91	0.61	0.30	0.92	0.51	-0.16	I	0.13	0.21	0.29	0.21	0.72	0.19	0.01	VII
0	1	-0.05	-0.89	0.63	0.32	0.94	0.53	-0.14	I	0.14	0.25	0.22	0.26	0.73	0.17	0.32	I
0	3	-0.53	-1.38	0.14	-0.17	0.45	0.04	-0.63	VI	0.13	0.28	0.15	0.21	0.46	0.17	0.22	I
0	11	-0.09	-0.94	0.58	0.27	0.89	0.48	-0.19	I	0.12	0.21	0.32	0.18	0.74	0.17	0.21	I
0	25	0.09	-0.75	0.77	0.45	1.08	0.67	-0.01	VII	0.14	0.17	0.47	0.24	0.71	0.23	0.25	I

Chapter 7

Discussion and interpretation

In this chapter, an overview of all steps and procedures is presented. Challenges arising from the inability to subgroup CLBP patients with satisfactory medical interpretation are considered again. Therefore, the focus is put on providing satisfactory subgrouping models that need to prove their discriminative power in classifying and subgrouping CLBP patients, alongside being interpretable from the medical point of view. General concepts driving contextual feature modeling are re-iterated again, where some of the relevant aspects contributing to LBP detection and differentiation are set to be: muscles coordination, co-activation, changes in muscle activity trends, and fatiguing processes.

As the first step in extracting meaningful myoelectric information, a range of raw (primary) features, mostly from the time domain, were examined. Based on the correlation and statistical analysis, six feature candidates were selected, namely, SSC, PE, WAMP, VAR, ZC, and RVD. In the next step, contextual (secondary) features were constructed where specific LBP-related domain knowledge was exploited. The resulting secondary feature groups (Coordination, Co-activation, Trends, Fatigue indices, Primary set) were compared among each other, as well as combined together (the Complete set) and validated through the classification task of differentiating HS and LBP groups. Satisfactory results were obtained, with the highest contributions coming from Coordination and Co-activation secondary feature groups, individually, whereas the overall best accuracy results were obtained by exploiting the Complete set.

As part of the further confirmation of the discriminative power of the proposed procedures and LBP modeling, a detailed classification modeling and groups differentiation analysis was performed. Namely, separate classification pairs for HS vs. LBP, HS vs. RLBP, HS vs. CLBP, and CLBP vs. RLBP with accuracy, precision, and sensitivity were cross-validated and analyzed. Respective classification success metrics showed very good results for HS vs. LBP and HS vs. RLBP pairs and fairly good results for HS vs. CLBP and CLBP vs. RLBP classification results. Alongside, the NCA feature selection technique was proved to be a very good and applicable technique for this type of classification and differentiation problem, especially

when a prerequisite of keeping the features in the original interpretable domain is immanent, which in this case with employing NCA was obeyed. As an additional insight, classification models based on SVM, *k*-NN and subspace ensemble techniques were proved to provide the best classification results for respective LBP models.

As the next important step, the obtained results, based on NCA-derived models, were committed to subgrouping procedures with a focus on primarily finding clustering patterns for the CLBP patients group. For each CLBP patient, a respective CLBP profile, based on radial representation, was visually inspected and five initial candidate subgroups were isolated. This preliminary procedure was verified with a more formal procedure based on unsupervised clustering techniques. SOM and HCA were employed where it was shown that HCA provides consistent and interpretable clustering results that can be related to the initially selected subgroup candidates quite well. Hierarchical tree dendrogram additionally suggested some extra clusters (groups) to be included in our CLBP subgrouping landscape, thus a total of seven subgroups were taken into consideration: *Regular CLBP pattern*, *Inhibited pattern*, *Type A coordination imbalance pattern*, *Type B coordination imbalance pattern*, *Exceeding coordination pattern*, *Near-regular pattern*, and *Co-activation imbalance*. To enable single simple metric for assigning any CLBP patient profile to the best fitting subgroup, a new metric called *CLBP Pattern Distance* (CPD) was introduced, with minimum distance to an averaged subgroup profile template (Fig. 6.17), as a criterion (6.3).

The respective classification and clustering results were supplemented with the corresponding medical interpretation taking also into account research insights from other authors and studies and correlating them with the newest insights presented in this PhD thesis, suggesting that LBP characteristics expressed via myoelectric activity can quite differ among the subjects (profiles) and uniform "one-size fits all" approach is evident not to hold. This concept introduced here now enables a more personalized approach when required to better understand the background of LBP-driven neuromotor control and impairments, that result in the individualized coordination, co-activation, and compensation adaptations emerging to avoid pain and enable painless movements. Such deeper insights support the aim of installing more individualized medical treatments with efforts to strive for better and more predictable rehabilitation outcomes.

7.1 About feature modeling

The focus of this study was to model such feature sets that enable satisfactory detection of LBP patients, alongside a foundation to further enable subgrouping within heterogenous LBP groups and with the ability to interpret results from a medical point of interest. Features were constructed with a motivation to reflect certain neuromuscular and motor control changes per-

tained to LBP, as discussed, and elaborated in a number of studies [7, 12, 27, 29, 31, 176, 177, 178, 179]. Measures of LBP coordination, co-activation, muscle fatigue, and trends tracked throughout the isometric exercise duration were selected as the key indicators for LBP classification modeling in this research.

The coordination aspect was considered as different studies were reporting asymmetric bilateral (left-right) muscle activation patterns to be observed for LBP patients, relating to motor control impairments, in the lumbar region during functional tasks [97, 180], compared to healthy subjects where such lumbar muscle coordination exhibited more symmetric patterns. Additionally, Liu *et al.* proposed modeling the activity coordination network between lumbar muscles based on flexion-extension tasks performed in sEMG electrode matrix setup. Results demonstrated that healthy subjects clearly exhibited globally symmetric patterns between the left and right side of sEMG channels, whereas the LBP patients group showed the loss of global symmetric patterns. However, Revees *et al.* [181] compared varsity athletes with and without low-back injury history and reported no group differences in the imbalance between sides, during an isometric trunk extension exercise. Some earlier studies by Roy *et al.* [31] also reported that left-right differences were present in both LBP and control groups, recorded during isometric contraction exercise, thus not being specific to one group particularly. The presented results demonstrated certain ambiguities in reported findings, likely to come due to different setups as well (e.g., isometric exercise or functional tasks, athletes, or the general population, few sEMG channels or electrode arrays, etc.), with no univocal conclusion about how the coordination patterns contribute to the groups' differentiation. Hence, the idea to include coordination aspects into consideration and further investigate such aspects of muscle activation imbalances between sides, seemed plausible.

In this PhD thesis, primarily the bilateral relations for ULES and LLES were analyzed (section 4.1), thus not considering contralateral imbalances as in some other studies [177]. Moreover, our proposed coordination measure was based on similarity measures calculated intra-subject between left and right muscle sites, by pairing time-aligned time epochs ($L = 1000$ ms, $s = 50$ ms) between muscle sites throughout the whole exercise duration. This way, imbalances and differences were captured on an immediate localized scope. This approach is different from studies where measures (e.g., MDF, RMS) were calculated for each muscle site separately, and afterwards, such single valued measures were compared to establish a macroscopic significance of existing imbalance between sides [31, 177, 181].

To capture more subtle co-activation patterns across all four muscle sites that were examined, an additional measure of co-activation trigger score was introduced (section 4.2). When considering movement-related LBP differentiation based on the functional type of tasks (e.g., flexion-extension) it is shown that there is a significant difference in muscle activation patterns in the lumbar region between healthy and LBP groups [12, 178], like differences pertaining to

FP and AEP patterns with motor control impairments considered to be the mechanism for maintaining CLBP. Also, altered temporal patterns of co-activation [176] and activity (re)distribution [52] between muscle sites in the lumbar region can be observed in LBP patients, compared to healthy subjects, implying different activity alignments for otherwise highly synergistic behavior of back muscles [50]. In that course, an interest to capture such temporal activation events on a finer time-localized scope was raised. Furthermore, such observations are inherently easier to capture for the functional type of exercises performed, due to distinctive time-related events (e.g., sitting vs. standing, flexion vs. extension, etc.) that are analyzed and compared in the course of the time. However, for isometric exercise, as in the case of this research, a new measure (score) was required to capture different time-related events within and between muscle sites (i.e., sEMG channels) (Fig. 4.3), without an obvious external event-based trigger that would induce time-related myoelectric changes. For that reason, "alignment" and "misalignment" scores, for the respective primary feature in behind, were calculated. In this case, "alignment" assumed time-aligned co-activation (synergistic activation) across the majority of muscle sites, whereas for "misalignment", only individual muscle (single muscle site) activations were observed, i.e., no time-aligned co-activation across other muscle sites was observed.

To support further the construction of a new comprehensive model, fatigue indices and trends were included in the secondary features groups as well. Fatigue assessment, based on MDF linear regression, is a common concept utilized in the analysis of LBP, although not always straightforward results were reported. Namely, some papers reported higher values for the initial median frequency, f_0 (4.13) in LBP patients with statistical significance [17], whereas in some papers no significant difference between healthy and individuals with LBP was reported [40]. For the linear regression slope, k_{sl} (4.13), the reported results were also not univocally straightforward, however, most of the papers report statistically more significant results suggesting higher fatigability of muscles in the lumbar region for individuals with LBP [31, 54, 92]. To expand features set in a search of potentially useful LBP indices, beyond MDF and linear regression only, the trends feature group was introduced as a set of measures setting the relations for both intra-channel and between channels (bilaterally and ipsilaterally), as presented in section 4.3.

7.2 How to select the initial feature set?

As the first step in the feature modeling process, fourteen primary features (Table 3.2) were chosen as a basis to find those best-performing ones in the context of LBP detection. The focus was on TD features as the most commonly employed feature group in sEMG-based classification [85, 87]. The choice for primary raw features was also inspired by the feature vectors proposed by Hudgin *et al.* (MAV, WL, ZC, SSC) [123] and Du *et al.* (IEMG, VAR, WL, ZC,

SSC, WAMP) [124] employed in sEMG-based real-time HMI applications. Alongside, some of the TD features were a common choice for many LBP-related studies, analyzing changes in EMG amplitude for activation detection or exhibited force levels via RMS, VAR, MAV or IEMG [9, 18, 21, 23]. The initial set of primary (raw) features employed was further populated with KURT and SKEW, as higher-order statistics measures analyzed in [84, 88], MDF as used in a number of LBP related studies [17, 18, 19, 20, 22, 23], and PE as a measure of non-linear signal complexity [126]. Finally, RVD, as a newly constructed feature, was added to the feature set with the intention to detect local changes in sEMG due to force level changes, activation bursts, or co-activation triggers. Such a feature was not found in any other analyzed research or review.

Individual classification results (Table 5.1) and corresponding feature ranks (Table 5.2) indicated that the best performing features (SSC, PE, WAMP, VAR, ZC, RVD) correlated more with Du's feature vector (IEMG, VAR, WL, ZC, SSC, WAMP) than with a Hudgin's feature vector (MAV, WL, ZC, SSC), though, the latter consisting of only four vector elements. In our study, IEMG, WL, and MAV demonstrated below-average results, which was a significant difference compared to those features selected for Du's and Hudgin's vectors. Furthermore, the MDF feature demonstrated below-average individual classification results, which coincides with observations reported by Phinyomark *et al.* [85], stating that EMG features calculated in the frequency domain do not provide good results in EMG signal classification. However, interestingly, the SSC feature, characterized by carrying certain frequency relation information, is shown to be one of the best features in sEMG-based classifications, namely, being part of Hudgin's and Du's feature vectors, as well as best ranking individual feature in this study.

Alongside, correlation analysis confirmed high correlation and redundancy among subset of features (MAV, IEMG, LD, RMS, VAR) (Fig. 5.2), where in this research best performing feature among these was VAR, unlike MAV in [85, 123], or RMS as employed in many other studies [18, 21, 23]. It is also noteworthy to state that no bias towards boosting best feature-classifier pairs was promoted, but results across all classifiers were taken into account for the selection of primary features subset that was further employed in constructing secondary feature groups.

To validate the best performing features presented in this PhD thesis, a comparison among different primary feature subsets, employed to construct the complete LBP detection model based on all secondary feature groups, was performed. The results, as per Table 5.3, showed feature sets (Top 5, Top 5+PE, Uncorrelated) composed of individually best performing and uncorrelated features outperformed Du's and Hudgin's features vectors. Such findings once more support the necessity of defining such feature sets and classification models that are suited for LBP-specific domains of problems, unlike some other domains, like hand motion or gesture classification [124] and prostheses control [123]). Interestingly, four elements Hudgin's feature

vector performed better than six members Du's vector. Individually, the "Uncorrelated" features set provided overall best accuracy results which can be explained by exploiting a full diversity of all uncorrelated features in defining the LBP classification model. However, to keep the model manageable in terms of interpretability, a feature set consisting of six primary features, SSC, PE, WAMP, VAR, ZC, RVD (namely, Top 5+PE feature set), resulting in twenty NCA components, was selected as optimal in terms of both accuracy and complexity.

7.3 Do contextual features contribute to LBP detection?

By analyzing the classification accuracy results for the secondary feature groups individually (Table 5.4), it can be observed that none of the feature groups excelled in detecting LBP patients alone. Best classification accuracy results, among created feature groups, were demonstrated for the coordination and co-activation measures with 0.91 (for fine *k*NN) and 0.88 (coarse SVM), respectively, where median accuracy results were 0.81 and 0.80, respectively. However, when different features groups were combined as input into the classifier models, a complete set of features demonstrated better results with higher accuracy (0.90, 0.91 for the fine resolution and subspace *k*NN, respectively, and 0.93, 0.94, 0.94 for the medium Gaussian, cubic and quadratic kernel SVM classifiers, respectively), as well as higher median accuracy (0.84) across all classifiers.

These results contribute to the idea of introducing more diversity into describing LBP through feature modeling, as it was hypothesized. Such concept was also following the questions and comments posed by Dieën *et al.* on differences in motor control between individuals with and without LBP [29] and whether such variations, in motor control issues, might contribute to more personalized approaches in LBP diagnostics and rehabilitation consequently [179], thus avoiding "one-size fits all" approach [8, 25].

Although features diversity inherently assumes introducing more complexity into the classification models, this is inevitable if an individualized approach is aimed to be established. This was evidenced by the number of NCA components required to describe our complete model for detecting subjects with LBP (twenty components employed) compared to coordination and co-activation models alone (with eleven and twelve NCA components, respectively). However, by introducing contextual and more meaningful features and measures, the portion of complexity can be lifted through more immediate medical interpretation. This is contrary compared to some other studies where a significant number of features was employed, but with more difficult to interpret contributions of respective features [23, 50]. This challenge especially holds when deep learning techniques are employed [13].

Moreover, different secondary feature subsets were observed across all cohort differentiation pairs as presented in Table 6.1, where Coordination, Co-activation, and Trends, as major feature

groups, were present in all models. This suggests that diverse secondary feature groups might indeed support favorably LBP detection and differentiation through the proposed measures for contextual modeling. Otherwise, classification and differentiation results, based solely on the primary feature group and fatigue indices, as seen in other studies, would have demonstrated somewhat more success itself, especially considering the relevant features pool given at hand (twenty-six different fatigue-based features and six primary features per each channel). In this PhD thesis, that was shown not to be the case, where moreover, coordination and co-activation characteristics were the key LBP modeling contributors.

7.4 Do differentiation models provide satisfactory results?

The proposed method and the respective modeling demonstrated to be successful in differentiating the LBP group (a group combined from CLBP and RLBP patients) against HS, with the accuracy of up to 0.94 and high sensitivity of up to 0.97 (Table 5.5). However, for such accurate results, the positive contribution of successfully detecting patients with RLBP is likely imminent. Namely, HS vs. RLBP classification pair provided median accuracy of 0.96, alongside high precision and sensitivity, with median values of 0.97 and 0.99 across all classifiers, respectively. Nonetheless, it is deemed fair to acknowledge that HS vs. LBP model had to account for a significant variety among groups and subjects at test, which also resulted in the largest number of NCA components (twenty), pointing to a complex model containing 4 out of 5 secondary feature groups (only MDF-based fatigue indices were not included in this particular model). On the other hand, differentiation between CLBP and HS groups only, was somewhat less successful, but still with a satisfactory result for the best classification results in the range of 0.88 to 0.90. Similarly, differentiation between CLBP and RLBP showed to be a challenging task, where fourteen NCA components were used to describe the model with at best classification accuracy of 0.86 and 0.89 achieved for subspace and fine kNN classifier, respectively. Furthermore, CLBP vs. RLBP model was the most diverse model including all of the secondary features group, thus pointing to motor control and neuromuscular complexity (or variability) behind. Overall, these results are in alignment with so far commented challenges in defining classification models for non-specific CLBP problem description, as well as related to subgrouping tasks, as indicated in the case of CLBP vs. RLBP.

Interestingly, in all of the classification pairs, time-to-endurance (T_{emg}), as a single non-sEMG parameter utilized in this study, did not contribute to any of the classification models. This parameter alone, was shown to be a good indicator for the presence of LBP in different studies [9, 31, 40], although sometimes with a contradictory result, as in work by Lariviere *et al.* where, for healthy subjects, shorter endurance was reported [51].

Additionally, it is also worth stating that the initial primary feature selection, resulting in

SSC, PE, WAMP, VAR, ZC, and RVD feature set, was defined based on the best fit for HS vs. LBP classification. Moreover, all classifiers were employed out-of-the-box with a predefined set of parameters. This is leaving room for further parameter optimizations that could lead to improved classification and differentiation results for more challenging tasks (namely HS vs. CLBP, or CLBP vs. RLBP).

When analyzing the classification results across twenty-three classifiers and their variants exploited, it is shown that SVM and *k*NN-based classifiers were consistently providing the best classification accuracy results. More concretely, quadratic and cubic kernel SVM, together with fine resolution and ensemble subspace *k*NN, demonstrated to be the most successful in dealing with LBP detection and related differentiation tasks. On the other hand, classifiers based on discriminant analysis and decision trees (together with ensemble boosted and bagged trees) were the least successful. Below average classification success results for LDA were somewhat surprising as it was one of the most commonly employed classifiers in LBP classification related tasks [15, 16, 17, 18, 19]. Somewhat better results were demonstrated by QDA, as dealing more successfully with more complex classification tasks in the quadratic separation surface, however, often suffering from resulting singular covariance matrix calculations.

In this study, a certain caution was posed around *k*NN-based classifiers. Namely, the fine resolution *k*NN classifier was labeling each sample at test based on only one ($k = 1$) nearest neighbor. Such an approach is very prone to outliers and noise. Additionally, NCA-based feature selection is based on a 1-NN classifier, thus certain bias towards respective *k*NN classification models might have been induced. To assess these effects on concrete LBP classification tasks, further investigation would be required. However, the obtained results in this study showed that SVM outperformed *k*NN classifiers, so it was deemed NCA not to have a decisive bias-prone impact on the classification success for the methods proposed. Moreover, for the ensemble subspace *k*NN models a medium resolution with $k = 10$ was employed.

7.5 Is subgrouping of CLBP patients possible?

Subgrouping of CLBP patients was posed as one of the major challenges within the domain of LBP research, where >90% of LBP patients are categorized as with non-specific chronic LBP, in real clinical practice. By continuing on the results and models obtained by two-class HS vs. CLBP classification task (Table 5.5 and Table 6.1), it was possible to construct CLBP profiles (Fig. 6.2) that were visually inspected and grouped. Even this basic manual procedure enabled the detection of five groups, that were deemed distinctive, within the CLBP data set at hand (Fig. 6.3), where overall 15 out of 29 subjects were assigned to one of the five subgroups. Alongside, four additional profiles (subjects) were annotated as candidates for one of the proposed subgroups.

This non-objective grouping method had to be elevated. For that reason, some of the well-known clustering techniques were employed, namely SOM and hierarchical clustering. Clustering results obtained by SOM were not contributing to the overall goal of establishing a consistent procedure with reproducible results by enabling inference of the exact connections between each individual CLBP subject and the respective subgroup that the respective subject was assigned to. With SOM, this expectation was not met. On the other hand, by employing HCA and dendrogram visualization (Fig. 6.11) this target was achieved. Namely, a link between the CLBP subject and the respective placement in the hierarchical clustering tree was enabled. Thus, HCA was exploited to verify whether the initially proposed grouping could be confirmed with a more formal mathematical procedure. With this approach, it was confirmed that all of the initial 15 subjects were consistently grouped in a similar way as provided by HCA. Alongside, 3 out of 4 candidate profiles were also initially assigned a proper cluster (i.e., subgroup). Furthermore, by exploiting HCA, 2 additional potential subgroups were detected and added to the existing set of five subgroups, thus resulting in the following seven subgroups:

- *Regular CLBP pattern*
- *Inhibited pattern*
- *Type A coordination imbalance pattern*
- *Type B coordination imbalance pattern*
- *Exceeding coordination pattern*
- *Near-regular pattern*
- *Co-activation imbalance pattern*

Alongside, existing subgroups were populated with the neighboring observations, thus filling in the endpoints of the hierarchical tree. As a result, a total of 25 subjects, out of 29, were grouped into seven subgroups that were proved to be distinctive by following the formal mathematical procedures as well. Also, four subjects were left with no subgroup assignment as deemed not to directly fit into any of the proposed subgroups, thus continued to be considered non-specific.

Alongside sEMG data, supplementary metadata information was also examined. Namely, BMI, exercise duration (time-to-fatigue), ODQ, RDQ, VAS, and FFK (left and right) values were considered to verify how the potential subgroupings, only based on these parameters, resonate with the earlier proposed subgroupings derived from visual inspection and HCA. Expectedly, it was shown that metadata-based subgrouping was not providing clear grouping patterns that could be more closely correlated with our subgroupings (Fig. 6.16), thus confirming that such metadata themselves are not sufficient for LBP inference and patient's subgrouping. In that context, it is once more confirmed that sEMG-based analysis provides additional rich information that enables a deeper LBP inference based on myoelectric activity with respective neuromotor control processes behind.

To round up the proposed procedure, a new distance metric called *CLBP Pattern Distance* (CPD) was introduced (6.2). The motivation was to provide a simple mathematical procedure and metric that will enable any new CLBP patients to be assigned to any of the existing subgroups (or left out further as non-specific). To test the given procedure, all of the existing subjects' profiles were submitted to this task of subgroups re-assignment, by exploiting the given metric against the constructed profile templates, one for each of the seven groups (Fig. 6.17). The outcome reconfirmed the same assignment results as the ones obtained in the previous steps by the combined visual+HCA subgrouping procedure (Table 6.4).

The overall obtained results provided a strong notion that subgrouping within the CLBP group, into more homogeneous groups, is possible but requires a suitable multifactorial model with the respective set of components. Such a model should be able to capture deeper and more complex myoelectric characteristics, as described via here proposed features that are most likely able to reflect the underlying neuromotor and biomechanical processes.

7.6 Can we interpret the proposed models?

Interpretability has been posed as another goal in this PhD thesis, following the satisfactory classification results and meaningful clustering suggestions. As emphasized in several places, feature selection was based on NCA, thus, the features were preserved in the original domain. This, combined with the approach of introducing the contextual feature modeling, has opened the opportunity to introduce also more meaningful explanations for the respective classification and clustering models. Consequently, such a method enabled more insights into the individualized sEMG-expressed characteristics related to LBP.

Looking at the example of HS and RLBP differentiation, it can be observed that RLBP can be differentiated from HS with very high accuracy (0.96 median value) by only taking three feature components (Table 6.1). Moreover, such a strong classification distinction can be translated into a discriminative visual representation, as shown in Fig. 6.1, with a very clear separation of grouping between RLBP and HS. This was confirmed with the Wilcoxon rank sum test at 5% significance level. Furthermore, such visualization now provides a direct insight into the variations of the underlying RLBP characteristics (when compared with the HS group). It can be observed that these contextual characteristics are dominantly related to the LLES region with coordination (bilateral) and co-activation (synergistic time-alignment) aspects. Namely, the provided information is suggesting that patients with RLBP are dominantly exhibiting reduced bilateral coordination and co-activation in the LLES region compared to HS individually. Also, looking into the SSC-based feature (Tr-2) with a slower decreasing frequency of activation triggers trend for RLBP patients, one of the interpretations could be that, in RLBP patients, at least one side of ULES, has significantly dominant activation of fast-twitching muscle fibers

(type II), compared to healthy subjects. These interpretations follow some of the previous observations reported on asymmetrical functioning patterns in low back muscles in patients having radiculopathy [22], where the reported asymmetry is likely to be additionally emphasized by the characteristic one-sided radiating pain.

On the other hand, differentiation between HS and CLBP introduced more components (thirteen) into the model, thus leading to an assumption of a more complex, i.e. more diverse, neuromuscular and motor control dynamics behind. This complexity and diversity can be captured directly from Fig. 6.17 where the averaged CLBP profiles templates for seven distinctive CLBP patterns are shown. In the background, it can be observed that no clear separation between HS and CLBP groups (looking from the overall groups' levels) is possible, moreover indicating quite heterogeneous CLBP profiles within the CLBP group itself (Fig. 6.15), unlike the more homogeneous RLBP case (Fig. 6.1). However, this again concords with so far discussions on multifold variations in LBP patients and the necessity of subgrouping to tackle non-specific CLBP [25, 29].

One subset of exploited NCA features, as part of the HS vs. CLBP differentiation model, is based on the PE feature (section 3.2.2.12). This raw feature is characterized by the ability to quantify the level of nonlinearity in the signal and certain complexity behind. In the context of myoelectric activity recorded and analyzed, in this PhD thesis, it is assumed that such feature and the respective quantifiable measures (namely, NCA components 5, 7, and 13 as listed in Table 6.2) provide insights into more advanced neural motor control compensation and redistribution strategies that result in more complex coordination and co-activation patterns. Complexity can be expressed as a superposition of different motor control components or strategies, namely activation or deactivation of certain motor units, engaging additional motor units in the same region, or redistributing the load onto synergistic muscles and among different muscle fiber types (namely type I and type II).

To provide deeper insights into the nature of CLBP subgroup patterns and, presumably, to introduce meaningful medical interpretation into this concept, each of the respective seven proposed subgroups is discussed in more detail in the following sections. Such interpretations, supported by classification and clustering results, might further contribute to the insights into the inconsistencies reported in the number of studies that result in contradictory interpretations. Elaborated commentary on these challenges were also given by Dieën *et al.* [9, 29].

7.6.1 Regular CLBP pattern

Regular pattern is set of profiles that most closely follow the median CLBP profile and are considered to represent a typical CLBP patients (Fig. 6.4). From a statistical point of view, this makes sense but does not necessarily need to hold. Namely, in cases of multimodal distributions, the median value can significantly differ from the majority of the population, where the segment

around the median value can represent only a small portion of the overall population. In this case, seven profiles were categorized to correspond to the regular pattern, thus individually being the single largest subgroup (∼25%) within the CLBP group at hand and consequently being considered the central (cornerstone group) within the CLBP subgrouping landscape.

This subgroup is characterized by less decrease (or even increase) in the frequency of activation triggers (based on SSC), both in LLES and ULES, compared to HS, indicating potentially more engagement of fast-twitching fibers (type II). Further, there is an indication of the similar complexity of activity patterns between the left and right sides in LLES, assumably more bursty, somewhat less predictable compared to HS, based on PE, further pointing to potentially more engagement of fast-twitching fibers as those resulting in more irregular myoelectric patterns (combined with type I) than type I fibers only. In the LLES region, a portion of more individual activation triggers, compared to HS, are exhibited, based on WAMP. However, in ULES region, more pronounced co-activation, compared to HS, is observed, pointing to presumably more fast-twitching contraction in this region.

The proposed interpretation view might point to assumably more myoelectric activity in ULES regions and thus colliding with some general observations in other works, where, for the CLBP group, muscle activity in LLES is reported to be more pronounced than in ULES area [9, 30, 31]. However, the presented view is not explicitly stating that muscle activity in the LLES area is less pronounced than in the ULES area. It only points out that apparently compensation mechanisms in the LLES areas were not sufficient, thus expanded redistribution and compensation strategies, involving the ULES area, were employed.

If looking into metadata, certainly high values for ODQ, RDQ, and VAS can be observed, supporting the fact that CLBP impairments recorded through sEMG are also reflected in the subjective sensation of pain and LBP-caused disability. Even more, for these seven subjects at hand, the average endurance time of ∼80 s is noticeably below the average (∼105 s) for the whole CLBP group at hand (29 subjects).

7.6.2 Inhibited CLBP pattern

When CLBP patients' profiles in Fig. 6.5 are analyzed, it can be observed that all coordination, co-activation, muscle activity, and triggering aspects are significantly inhibited compared to median values for both the HS group and CLBP group.

This pattern is characterized by the significant co-activation impairments in both LLES and ULES regions. There is also a strong indicator (based on SSC) of most likely no change in the frequency of activation triggers, compared to both HS and the remaining CLBP patients, exhausting fast-twitching fiber's force and endurance capacity supported by the fact that for this pattern subgroup the exercise duration times were on average <50 s, least among all other subgroups. One of the possible interpretations for a such distinctive set of profiles is that pa-

tients adapt to pain due to musculoskeletal deficiencies in manifold ways. In some cases, these adaptations lead to avoidance-like inhibited behavior as a strategy to avoid pain or protect from further damage or injury on the tissue level [27]. For such outcomes, the psychological and anticipatory pain avoidance effects, reflected through the neuromuscular coding, also play a role. These outcomes can be also related to the meta-information collected, where the highest BMI, relatively low FFK values, and relatively high ODI and RDQ scores were observed, indicating such subjects suffering from significant pain or disability, alongside reduced physical fitness state. Moreover, excessive pain is often accompanied by a psychological factor of fear [7], contributing to short exercise duration times, and preventing certain LBP indicators to develop fully in the course of sEMG recording.

7.6.3 Type A coordination imbalance CLBP pattern

For the CLBP patient's *Type A coordination imbalance* profiles shown in Fig. 6.6, significant muscle activity and changes in the LLES region can be observed, even exceeding statistically significant high values primarily characteristic for HS group (components 1, 5, 7, 9). Moreover, very distinctive, and somewhat opposite, expressions on LLES left-right coordination were observed. More concretely, significant coordination pertained to the levels of muscle activity patterns complexity (component 7, PE) can be observed. This coordination, in terms of intensity, even exceeds such type of neural coding that is inherently more characteristic for healthy subjects, even with statistical significance when comparing HS with the overall CLBP patient groups. On the other hand, a significantly inhibited level of LLES left-right force balance, even when compared to the rest of the CLBP group itself, was observed. Such inhibition points to a potentially reduced ability to maintain the same exerted force levels on both sides. Additionally, apparently more time-decreasing frequently of activation triggers is observed (components 1 and 2, based on SSC), assumably leading to more stable (endurable) contractions in both LLES and ULES regions. Such observation is correlated with the fact that subjects in this group exhibited long exercise durations ($\sim$150 s on average), which is even higher than the average endurance times for the HS group ($\sim$140 on average), moreover demonstrating high FFK and relatively low ODQ, RDQ and VAS values, quite comparable to HS. Another expression noticed is related to the increased co-activation both in ULES and LLES regions, primarily on the left side. Here, for LLES left side, excessive muscle activity was detected, resulting in lots of hits and misses in co-activation triggering (components 9 and 10), where, overall, also more synergistic misalignments were detected (components 10 and 12), compared to the HS group.

Possible interpretation for the adaptation and compensation expressions for this subgroup is proposed as follows: certain motor control impairments or tissue deficiencies in the musculoskeletal system were (over)compensated in such a way that the central nervous system engaged such neuromotor control to compensate the observed coordination and force level im-

pairments. This consequently resulted in an increased underlying (co-)activation in the LLES region as well as likely engaging more slow-twitching muscle fibers in both ULES and LLES regions with a goal to endure the trunk extension exercise as long as possible (which was successfully achieved with average endurance times $\sim$150 s). It could be also concluded that subjects in this group were not suffering from psychological "tiredness" or fear of injury, but had the physical and psychological capacity to compensate for the deficiencies. The imbalances and compensations, characteristic for this subgroup, are quite distinctive when compared to both HS and the remaining CLBP group, where for the HS group in general, the mentioned force-wise and complexity-wise coordination aspects are much more in balance on the individual subjects' levels.

7.6.4 Type B coordination imbalance CLBP pattern

Type B coordination imbalance pattern (Fig. 6.7) is shown to be another distinctive CLBP subgroup pattern. Somewhat opposite to *Type A coordination imbalance pattern*, this pattern shows significantly reduced, or impaired, coordination level of muscle activity patterns complexity between the left and right sides in the LLES region. This presumably led to less bursty (more predictable) activation patterns taking more one-sided end of the exercise, resulting in left-to-right activation pattern impairments. On the other hand, regular CLBP levels of coordinated exerted force, that are comparable to the average levels of HS group, are observed in the LLES region. Additionally, a certain compensation in the ULES region is observed with the higher alignment of the increased frequency of (co-)activation triggers, supported by both left and right sides, compared to HS. In terms of the available metadata, although having only 2 subjects in this group, two distinctive characteristics can be observed, firstly, quite long endurance times (one 140 s, the other 177 s), and secondly, mid-range values for FFK (65°-70°), alongside with the highest ODQ, RDQ, VAS values, among all CLBP subgroups analyzed.

Possible interpretation is: subjects' overall physical fitness is evident to be on the above average level enabling long exercise endurance, despite the certain CLBP impairments in LLES region and a registered subjective feeling of pain and LBP caused disability. A such (subjective) sensation of pain did not impact the neuromotor control in a way to lead to an avoidance-like inhibited behavior (as for the *Inhibited pattern* discussed in section 7.6.2), but neuromotor redistribution and activation managed to compensate imbalances in LLES region side with more endurable (slow-twitching) muscle activation in the ULES region, where, combined with high fitness level, it resulted with sufficient force levels for sustaining long-endurance contractions.

7.6.5 Exceeding coordination CLBP pattern

Exceeding coordination is yet another quite distinctive CLBP subgroup pattern (Fig. 6.8). It is characterized by significantly above average values for force levels coordination in the LLES region, alonside including significantly reduced frequency of activation triggers on both left and right sides in LLES region, accompanied by presumably slower activation muscle contractions (type I). These coordination levels are moreover significantly higher compared even to the HS group. However, by the end of a contraction, a certain impairment in left-right relation in LLES regions for both activation triggers and force levels are observed. However, with no evident (co-)activation synchronization in ULES region supporting LLES, resulting in overall average CLBP endurance times. Metadata is characterized by the above average FFK values (comparable to HS), and low RDQ.

Possible interpretation for this profile pattern is: subjects pertained to this subgroup are deemed to engage in extra coordination and compensation mechanisms in LLES region in order to maintain the trunk extension contraction with presumably more slow-twitching muscle contractions. However, with no synchronization support from ULES regions, the overall capacities were limited (considering also subjects' physical fitness state, e.g., higher BMI observed or due to LBP disability or pain sensation perceived), resulting in the average CLBP endurance times.

7.6.6 Near-regular CLBP pattern

Near-regular pattern (Fig. 6.12) was isolated and added to set of subgroup patterns, as a result of HCA and dendrogram inspection (Fig. 6.14). It possesses certain similarities with the *Regular CLBP pattern*, thus the given name. Moreover, as per dendrogram hierarchy visualization, the CLBP patients' profiles, corresponding to this pattern, are sitting in the same section of the hierarchical tree (meaning, they are part of the same parent cluster). Certain characteristics pertained to this pattern are relating to more (co)activation observed in the ULES region, with more decrease in frequency of activation triggers (somewhat more slow-twitching fibers taking role compared to *Regular CLBP pattern*). Overall, more activation observed in the LLES region eventually resulting in somewhat more left-to-right difference end of exercise. Metadata show above average endurance times ($\sim$120 s) for the CLBP group, FFK ($\sim$80 °) values comparable to the HS group, but also relatively high ODQ and RDQ scores, comparable to the *Regular CLBP pattern*.

As for the possible interpretation: subjects in this group are fairly similar to regular CLBP patients, thus suffering from typical chronic LBP with certain adaptations to pain mechanisms being installed, where somewhat above average exercise durations are recorded (with subjects even not necessarily being in good physical fitness, as slightly above average BMI values were reported). This could be attributed to a certain level of LLES load being redistributed and

compensated where more synchronized ULES muscles activity overtook the load partially.

7.6.7 Co-activation imbalance CLBP pattern

Co-activation imbalance pattern (Fig. 6.13), as taking a smaller portion of the CLBP cohort at hand, is characterized by two somewhat more pronounced deviations from what is considered a regular CLBP patients' pattern according to subgrouping concept presented in this PhD thesis. Namely, less coordinated levels of exerted muscle force are observed in the LLES region compared to the median CLBP profile. This is accompanied by more pronounced individual one-sided (left) LLES muscle activations compared to *Regular CLBP pattern* and HS, alongside with an increased frequency of activation triggers on the ULES left side. Altogether, this resulted in the second least endurance times. As per metadata, this subgroup is also characterized by the lowest average FFK values (comparable to RLBP group) and among the highest ODQ, RDQ, and VAS values.

Possible explanation could go in direction of: significant impairments in LLES and ULES regions were not successfully compensated by either coordination mechanisms between left and right side (LLES), nor by the more dispersed redistribution and co-activation strategies between LLES and ULES regions, although certain compensation efforts have been seen on the both left LLES and ULES. Overall, the metadata relating to pain and disability aspects, together with short endurance times, suggest that pain and musculoskeletal deficiencies have prevailed over the compensation and adaptation mechanisms, where certain avoidance-like aspects neuromotor control could have also been assumed.

Chapter 8

Contributions and conclusion

In this chapter, an overview of overall contributions, as well as novel or non-standard approaches or methods utilized, is given, such that have not been earlier utilized or detected in other works relating to sEMG-based LBP analysis. As a final step, we provide a summary with key scientific contributions as part of this PhD thesis.

For start, a novel approach for LBP patients' detection and differentiation with contextual feature modeling has been proposed. Contextual LBP characteristics were modeled via secondary feature groups (Coordination, Co-activation, Trends, and Fatigue indices) derived from the simple raw (primary) features. NCA as the feature selection method was employed in order to preserve the features in the original domain. Therefore, a basis for further LBP patients subgrouping has been established by providing a closer insight into the LBP-related neuromuscular and motor control characteristics. Additionally, new visualization of discriminating components via patient profiles has been proposed, thus facilitating medical interpretations.

It is noteworthy emphasizing again that the minimum measurement protocol setup was installed by purpose. Namely, the intention was to remove any supporting factors, that could have created any positive bias toward the end results and conclusions presented in this PhD thesis. Thus, the approach presented has avoided using multiple muscle sites and electrodes (only 2 pairs of bilateral electrodes in ULES and LLES regions were employed); it avoided using the dynamic (functional) type of exercises; in features models construction only sEMG-based features were used (endurance time was implicitly non-electric measure utilized); only out-of-the-box classification and clustering methods were employed, avoiding any parameters tuning or optimization. Such an approach has enabled us to fully concentrate on LBP features modeling and interpretation.

The initial (primary) TD feature set, consisting of six features SSC, PE, WAMP, VAR, ZC, and RVD, has been proposed, validated, and consequently used to create secondary feature groups. A novel RVD feature had been created and introduced in this PhD thesis to capture local myoelectric changes in muscle force or energy levels. By checking the NCA components

in Table 6.1, it can be observed that the features based on RVD took a significant part in all the resulting classification models, especially useful in co-activation triggers analysis. This suggests RVD to be a relevant feature choice and good discriminator in respective sEMG-based contextual LBP models. Similarly, PE is contributing to all models, except for HS vs. RLBP differentiation, as there a less complex model pertains.

Also, DTW and Spearman distance have been validated as similarity metrics for LBP coordination measures. These metrics have not been seen in other LBP studies, however, in this PhD thesis, they have contributed to the final results (unlike some other metrics, like Euclidean distance or MI), thus proving a certain level of their usability in LBP analysis for more complex modeling, especially for the most challenging CLBP vs. RLBP differentiation tasks. In this PhD thesis, an approach with calculating coordination similarity measures intra-subject directly, between muscle sites on time-aligned feature sequences accounting for local changes, was also presented.

A novel contextual measure for tracking the co-activation alignments, across all muscle sites examined, has been constructed and employed. Co-activation scores have shown to be, alongside coordination-based features, the most contributing features to the overall classification results and models. Additionally, a benefit of applying autocorrelation to co-activation score measures has been confirmed. This approach proved to be useful in additionally improving the discriminating power of selected features, for all classification pairs except for HS vs. RLBP. One possible explanation for its success is that the autocorrelation by its nature has emphasized periodic or regular patterns of muscle activations, whereas cases with irregular activation patterns were diminished, thus further separating away those samples with initially less mutually distinctive co-activation triggering (as assumably more present in HS vs. CLBP case). In HS vs. RLBP case, this distinction was apparently evident even without autocorrelation transformation, which seems likely possible due to significant motor control and myoelectric differences between HS and RLBP.

The performance of twenty-three classifiers and their variants have been validated and the best classifiers have been detected, namely, cubic and quadratic kernel SVM, fine resolution kNN, and ensemble subspace kNN classifier.

Very good classification results have been achieved for different classification pairs between subject groups, especially for the LBP detection (CLBP and RLBP against HS) and RLBP differentiation (from HS), consequently opening room for meaningful medical interpretation.

Additional insight, as one of the more important takeaways captured, as part of this PhD thesis, is that in most of the models created and validated, it was shown that the absolute values obtained for muscle force or frequency-based characteristics were not playing the key distinctive role. Key aspects enabling successful differentiation among different LBP expressions were primarily tied to the relative relationships among muscle sites, as well as relations within the

individual muscle itself, as the ones relating to different phases of LBP emergencies. This was clearly observed for the differentiation tasks between HS and CLBP, as well as between HS and RLBP groups. On the other hand, some contribution of absolute feature values was observed for CLBP and RLBP differentiation only.

As a result of the subgrouping procedure, 25 out of 29 subjects were categorized into the respective subgroup with some level of certainty. Although small sizes of CLBP subgroups were tackled, some general distribution observations can be drawn. Seven subjects in the *Regular CLBP pattern* and four subjects in the *Near-regular pattern* subgroup, thus 11 out 29 ($\sim$35-40%) subjects could have been assumed to be part of a regular pattern in a broader sense (part of the same segment of hierarchical clustering tree). The second largest consolidated group is the one for *Inhibited pattern* where 5 subjects are assigned, thus representing $\sim$15-20%. The next most consolidated subgroup was the one for the *Type A coordination imbalance* with four subjects assigned ($\sim$15%). The remaining three subgroups (*Type B coordination imbalance pattern*, *Exceeding coordination pattern*, and *Co-activation imbalance*) had only 2 representatives per subgroup, thus <10%.

Based on the subgrouping results and created CLBP profiles, the NCA feature models at hand (derived for HS vs. CLBP differentiation), consisting of thirteen feature components, were analyzed in detail to provide meaningful and relevant medical interpretation. Alongside, certain relations were observed between such models providing subgrouping, on one side, and metadata, on the other side.

8.1 Key contributions

Here, key contributions in this PhD thesis are presented, and can be summarized with the following three tasks:

- **Classification task**: Enable successful differentiation between the LBP patients (CLBP and RLBP group) and control group (HS group)
- **Clustering task**: Enable meaningful subgrouping of CLBP patients
- **Interpretation task**: Provide meaningful interpretation based on the obtained LBP features model and the relevant medical interpretation behind

These challenging tasks were tackled in depth where satisfactory results were obtained with high classification accuracy results, for four distinctive differentiation classification tasks (up to 0.94 for HS vs. CLBP, up to 0.98 for HS vs. RLBP, up to 0.90 for HS vs. CLBP, and up to 0.89 for CLBP vs. RLBP). These relevant results confirmed that the proposed feature models have potential and sufficient discriminative power to be verified against the clustering task. The clustering procedure resulted in seven distinctive enough subgroups that could have been even isolated by visual inspection (by exploiting a certain level of domain expert knowledge). These

subgroups were confirmed with an out-of-the-box hierarchical clustering procedure with high alignment resulting between these two procedures (visual and HCA). In the end, the interpretation and description for each subgroup were provided, where the model-based insights were combined with the available metadata to promote a comprehensive approach toward medical-based interpretation. The overall intention of this PhD thesis was to contribute further to removing ambiguities in explaining the neuromuscular processes pertained to the CLBP group and providing more insight into possibilities to establish more consistent LBP diagnostics procedures as well as designing such medical treatment methods that would lead to more successful rehabilitation outcomes.

Bibliography

[1]James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A. *et al.*, "Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017", Lancet, Vol. 392, No. 10159, 2018, pp. 1789–1858.

[2]Anderson, G., "The epidemiology of spinal disorders", The adult spine: principles and practice, 1997.

[3]Breivik, H., Collett, B., Ventafridda, V., Cohen, R., Gallacher, D., "Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment", European Journal of Pain, Vol. 10, No. 4, 2006, pp. 287–333.

[4]Dieleman, J. L., Baral, R., Birger, M., Bui, A. L., Bulchis, A., Chapin, A., Hamavid, H., Horst, C., Johnson, E. K., Joseph, J. *et al.*, "Us spending on personal health care and public health, 1996-2013", JAMA, Vol. 316, No. 24, 2016, pp. 2627–2646.

[5]Dagenais, S., Caro, J., Haldeman, S., "A systematic review of low back pain cost of illness studies in the United States and internationally", The Spine Journal, Vol. 8, No. 1, 2008, pp. 8–20.

[6]Bardin, L. D., King, P., Maher, C. G., "Diagnostic triage for low back pain: a practical approach for primary care", Medical journal of Australia, Vol. 206, No. 6, 2017, pp. 268–273.

[7]Hodges, P. W., "Pain and motor control: from the laboratory to rehabilitation", Journal of Electromyography and Kinesiology, Vol. 21, No. 2, 2011, pp. 220–228.

[8]Rabey, M., Smith, A., Kent, P., Beales, D., Slater, H., O'Sullivan, P., "Chronic low back pain is highly individualised: patterns of classification across three unidimensional subgrouping analyses", Scandinavian Journal of Pain, Vol. 19, No. 4, 2019, pp. 743–753.

[9]van Dieën, J. H., Selen, L. P., Cholewicki, J., "Trunk muscle activation in low-back pain patients, an analysis of the literature", Journal of Electromyography and Kinesiology, Vol. 13, No. 4, 2003, pp. 333–351.

[10]Bazrgari, B., Xia, T., "Application of advanced biomechanical methods in studying low back pain–recent development in estimation of lower back loads and large-array surface electromyography and findings", Journal of Pain Research, Vol. 10, 2017, p. 1677.

[11]Tagliaferri, S. D., Angelova, M., Zhao, X., Owen, P. J., Miller, C. T., Wilkin, T., Belavy, D. L., "Artificial intelligence to improve back pain outcomes and lessons learnt from clinical classification approaches: three systematic reviews", NPJ Digital Medicine, Vol. 3, No. 1, 2020, pp. 1–16.

[12]Dankaerts, W., O'sullivan, P., Burnett, A., Straker, L., Davey, P., Gupta, R., "Discriminating healthy controls and two clinical subgroups of nonspecific chronic low back pain patients using trunk muscle activation and lumbosacral kinematics of postures and movements: a statistical classification model", Spine, Vol. 34, No. 15, 2009, pp. 1610–1618.

[13]Wang, N., Zhang, Z., Xiao, J., Cui, L., "Deeplap: a deep learning based non-specific low back pain symptomatic muscles recognition system", in 2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). IEEE, 2019, pp. 1–9.

[14]Moniri, A., Terracina, D., Rodriguez-Manzano, J., Strutton, P. H., Georgiou, P., "Real-time forecasting of sEMG features for trunk muscle fatigue using machine learning", IEEE Transactions on Biomedical Engineering, Vol. 68, No. 2, 2020, pp. 718–727.

[15]Roy, S. H., De Luca, C. J., Emley, M., Buijs, R. J., "Spectral electromyographic assessment of back muscles in patients with low back pain undergoing rehabilitation", Spine, Vol. 20, 1995, pp. 38–38.

[16]Buijs, R. J., Levins, J.-A., Newcombe, D. S., Jabre, J. F., "Classification of back muscle impairment based on the surface electromyographic signal", Development, Vol. 34, No. 4, 1997, pp. 405–414.

[17]Peach, J. P., McGill, S. M., "Classification of low back pain with the use of spectral electromyogram parameters", Spine, Vol. 23, No. 10, 1998, pp. 1117–1123.

[18]Reger, S. I., Shah, A., Adams, T. C., Endredi, J., Ranganathan, V., Yue, G. H., Sahgal, V., Finneran, M. T., "Classification of large array surface myoelectric potentials from subjects with and without low back pain", Journal of Electromyography and Kinesiology, Vol. 16, No. 4, 2006, pp. 392–401.

[19]Candotti, C. T., Loss, J. F., Pressi, A. M. S., de Souza Castro, F. A., La Torre, M., de Oliveira Melo, M., Araújo, L. D., Pasini, M., "Electromyography for assessment of pain in low back muscles", Physical Therapy, Vol. 88, No. 9, 2008, pp. 1061–1067.

[20]Caza-Szoka, M., Massicotte, D., Nougarou, F., Descarreaux, M., "Surrogate analysis of fractal dimensions from SEMG sensor array as a predictor of chronic low back pain", in Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the. IEEE, 2016, pp. 6409–6412.

[21]Jiang, N., Luk, K. D.-K., Hu, Y., "A machine learning-based surface electromyography topography evaluation for prognostic prediction of functional restoration rehabilitation in chronic low back pain", Spine, Vol. 42, No. 21, 2017, pp. 1635–1642.

[22]Ostoji ć, S., Peharec, S., Srhoj-Egekher, V., Cifrek, M., "Differentiating patients with radiculopathy from chronic low back pain patients by single surface EMG parameter", Automatika, Vol. 59, No. 3-4, 2018, pp. 400–407.

[23]Du, W., Omisore, O. M., Li, H., Ivanov, K., Han, S., Wang, L., "Recognition of chronic low back pain during lumbar spine movements based on surface electromyography signals", IEEE Access, Vol. 6, 2018, pp. 65 027–65 042.

[24]Leboeuf-Yde, C., Manniche, C., "Low back pain: time to get off the treadmill", Journal of Manipulative & Physiological Therapeutics, Vol. 24, No. 1, 2001, pp. 63–66.

[25]Foster, N. E., Hill, J. C., Hay, E. M., "Subgrouping patients with low back pain in primary care: are we getting any better at it?", Manual Therapy, Vol. 16, No. 1, 2011, pp. 3–8.

[26]Panjabi, M. M., "Clinical spinal instability and low back pain", Journal of Electromyography and Kinesiology, Vol. 13, No. 4, 2003, pp. 371–379.

[27]Hodges, P. W., Tucker, K., "Moving differently in pain: a new theory to explain the adaptation to pain", Pain, Vol. 152, No. 3, 2011, pp. S90–S98.

[28]Bogduk, N., McGuirk, B., Medical Management of Acute and Chronic Low Back Pain: An Evidence-based Approach. Elsevier Health Sciences, 2002, Vol. 13.

[29]Van Dieën, J. H., Reeves, N. P., Kawchuk, G., Van Dillen, L. R., Hodges, P. W., "Motor control changes in low back pain: divergence in presentations and mechanisms", Journal of Orthopaedic & Sports Physical Therapy, Vol. 49, No. 6, 2019, pp. 370–379.

[30]van der Hulst, M., Vollenbroek-Hutten, M. M., Rietman, J. S., Hermens, H. J., "Lumbar and abdominal muscle activity during walking in subjects with chronic low back pain:

support of the “guarding” hypothesis?”, Journal of Electromyography and Kinesiology, Vol. 20, No. 1, 2010, pp. 31–38.

[31]Roy, S. H., De, C. L., Casavant, D. A., “Lumbar muscle fatigue and chronic lower back pain.”, Spine, Vol. 14, No. 9, 1989, pp. 992–1001.

[32]Juniper, M., Le, T. K., Mladsi, D., “The epidemiology, economic burden, and pharmacological treatment of chronic low back pain in France, Germany, Italy, Spain and the UK: a literature-based review”, Expert Opinion on Pharmacotherapy, Vol. 10, No. 16, 2009, pp. 2581–2592.

[33]Wenig, C. M., Schmidt, C. O., Kohlmann, T., Schweikert, B., “Costs of back pain in Germany”, European Journal of Pain, Vol. 13, No. 3, 2009, pp. 280–286.

[34]Bigos, S., “Acute low back problems in adults. clinical practice guideline no 14”, Agency for Health Care Policy and Research, Public Health Service, 1994.

[35]Kalichman, L., Kim, D. H., Li, L., Guermazi, A., Hunter, D. J., “Computed tomography–evaluated features of spinal degeneration: prevalence, intercorrelation, and association with self-reported low back pain”, The Spine Journal, Vol. 10, No. 3, 2010, pp. 200–208.

[36]Teyhen, D. S., Gill, N. W., Whittaker, J. L., Henry, S. M., Hides, J. A., Hodges, P., “Rehabilitative ultrasound imaging of the abdominal muscles”, Journal of Prthopaedic & Sports Physical Therapy, Vol. 37, No. 8, 2007, pp. 450–466.

[37]Ung, H., Brown, J. E., Johnson, K. A., Younger, J., Hush, J., Mackey, S., “Multivariate classification of structural mri data detects chronic low back pain”, Cerebral Cortex, Vol. 24, No. 4, 2014, pp. 1037–1044.

[38]Vogt, L., Pfeifer, K., Portscher, M., Banzer, W., “Influences of nonspecific low back pain on three-dimensional lumbar spine kinematics in locomotion”, Spine, Vol. 26, No. 17, 2001, pp. 1910–1919.

[39]Ashouri, S., Abedi, M., Abdollahi, M., Manshadi, F. D., Parnianpour, M., Khalaf, K., “A novel approach to spinal 3-D kinematic assessment using inertial sensors: Towards effective quantitative evaluation of low back pain in clinical settings”, Computers in Biology and Medicine, Vol. 89, 2017, pp. 144–149.

[40]Kankaanpää, M., Taimela, S., Laaksonen, D., Hänninen, O., Airaksinen, O., “Back and hip extensor fatigability in chronic low back pain patients and controls”, Archives of Physical Medicine and Rehabilitation, Vol. 79, No. 4, 1998, pp. 412–417.

[41]Geisser, M. E., Ranavaya, M., Haig, A. J., Roth, R. S., Zucker, R., Ambroz, C., Caruso, M., "A meta-analytic review of surface electromyography among persons with low back pain and normal, healthy controls", The Journal of Pain, Vol. 6, No. 11, 2005, pp. 711–726.

[42]Kleine, B. U., van Dijk, J. P., Lapatki, B. G., Zwarts, M. J., Stegeman, D. F., "Using two-dimensional spatial information in decomposition of surface EMG signals", Journal of Electromyography and Kinesiology, Vol. 17, No. 5, 2007, pp. 535–548.

[43]Olugbade, T. A., Bianchi-Berthouze, N., Marquardt, N., Williams, A. C., "Pain level recognition using kinematics and muscle activity for physical rehabilitation in chronic pain", in 2015 International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE, 2015, pp. 243–249.

[44]Liew, B. X., Rugamer, D., De Nunzio, A. M., Falla, D., "Interpretable machine learning models for classifying low back pain status using functional physiological variables", European Spine Journal, Vol. 29, No. 8, 2020, pp. 1845–1859.

[45]Biering-SØrensen, F., "Physical measurements as risk indicators for low-back trouble over a one-year period.", Spine, Vol. 9, No. 2, 1984, pp. 106–119.

[46]Encyclopædia Britannica, "Muscles of the abdominal wall", available on: https://www.britannica.com/science/human-muscle-system#/media/1/1346474/121139 [Accessed on 2022 Aug. 05].

[47]Encyclopædia Britannica, "Muscles of the back; Human muscle system", available on: https://www.britannica.com/science/human-muscle-system#/media/1/1346474/121600 [Accessed on 2022 Aug. 05].

[48]Vecteezy, "Spine anatomy", available on: https://www.vecteezy.com/vector-art/1879122-human-spine-anatomy-on-white-background [Accessed on 2022 Aug. 05].

[49]Pitcher, M. J., Behm, D. G., MacKinnon, S. N., "Neuromuscular fatigue during a modified Biering-Sørensen test in subjects with and without low back pain", Journal of Sports Science & Medicine, Vol. 6, No. 4, 2007, p. 549.

[50]Lariviere, C., Arsenault, A., Gravel, D., Gagnon, D., Loisel, P., "Evaluation of measurement strategies to increase the reliability of EMG indices to assess back muscle fatigue and recovery", Journal of Electromyography and Kinesiology, Vol. 12, No. 2, 2002, pp. 91–102.

[51]Larivière, C., Arsenault, A. B., Gravel, D., Gagnon, D., Loisel, P., "Surface electromyography assessment of back muscle intrinsic properties", Journal of Electromyography and Kinesiology, Vol. 13, No. 4, 2003, pp. 305–318.

[52]Hu, Y., Siu, S. H., Mak, J. N., Luk, K. D., "Lumbar muscle electromyographic dynamic topography during flexion-extension", Journal of Electromyography and Kinesiology, Vol. 20, No. 2, 2010, pp. 246–255.

[53]Leinonen, V., Kankaanpää, M., Airaksinen, O., Hänninen, O., "Back and hip extensor activities during trunk flexion/extension: effects of low back pain and rehabilitation", Archives of Physical Medicine and Rehabilitation, Vol. 81, No. 1, 2000, pp. 32–37.

[54]Hammill, R. R., Beazell, J. R., Hart, J. M., "Neuromuscular consequences of low back pain and core dysfunction", Clinics in Sports Medicine, Vol. 27, No. 3, 2008, pp. 449–462.

[55]Richardson, C. A., Snijders, C. J., Hides, J. A., Damen, L., Pas, M. S., Storm, J., "The relation between the transversus abdominis muscles, sacroiliac joint mechanics, and low back pain", Spine, Vol. 27, No. 4, 2002, pp. 399–405.

[56]Cholewicki, J., Juluru, K., McGill, S. M., "Intra-abdominal pressure mechanism for stabilizing the lumbar spine", Journal of Biomechanics, Vol. 32, No. 1, 1999, pp. 13–17.

[57]Paquet, N., Malouin, F., Richards, C. L., "Hip-spine movement interaction and muscle activation patterns during sagittal trunk movements in low back pain patients.", Spine, Vol. 19, No. 5, 1994, pp. 596–603.

[58]Alexiev, A., "Some differences of the electromyographic erector spinae activity between normal subjects and low back pain patients during the generation of isometric trunk torque.", Electromyography and Clinical Neurophysiology, Vol. 34, No. 8, 1994, pp. 495–499.

[59]Cohen, J., "Statistical power analysis for the behavioral sciences. 2nd", 1988.

[60]Finneran, M. T., Mazanec, D., Marsolais, M. E., Marsolais, E. B., Pease, W. S., "Large-array surface electromyography in low back pain: a pilot study", Spine, Vol. 28, No. 13, 2003, pp. 1447–1454.

[61]Nussbaum, M. A., "Static and dynamic myoelectric measures of shoulder muscle fatigue during intermittent dynamic exertions of low to moderate intensity", European Journal of Applied Physiology, Vol. 85, No. 3-4, 2001, pp. 299–309.

[62]Luoto, S., Aalto, H., Taimela, S., Hurri, H., Pyykkö, I., Alaranta, H., "One-footed and externally disturbed two-footed postural control in patients with chronic low back pain and healthy control subjects: A controlled study with follow-up", Spine, Vol. 23, No. 19, 1998, pp. 2081–2089.

[63]Travell, J., Rinzler, S., Herman, M., "Pain and disability of the shoulder and arm: treatment by intramuscular infiltration with procaine hydrochloride", Journal of the American Medical Association, Vol. 120, No. 6, 1942, pp. 417–422.

[64]Roland, M., "A critical review of the evidence for a pain-spasm-pain cycle in spinal disorders", Clinical Biomechanics, Vol. 1, No. 2, 1986, pp. 102–109.

[65]Zedka, M., Prochazka, A., Knight, B., Gillard, D., Gauthier, M., "Voluntary and reflex control of human back muscles during induced pain", The Journal of Physiology, Vol. 520, No. 2, 1999, pp. 591–604.

[66]Kravitz, E., Moore, M., Glaros, A., "Paralumbar muscle activity in chronic low back pain.", Archives of Physical Medicine and Rehabilitation, Vol. 62, No. 4, 1981, pp. 172–176.

[67]Cram, J. R., Steger, J. C., "EMG scanning in the diagnosis of chronic pain", Biofeedback and Self-Regulation, Vol. 8, No. 2, 1983, pp. 229–241.

[68]Holroyd, K. A., Penzien, D. B., Hursey, K. G., Tobin, D. L., Rogers, L., Holm, J. E., Marcille, P. J., Hall, J. R., Chila, A. G., "Change mechanisms in EMG biofeedback training: cognitive changes underlying improvements in tension headache.", Journal of Consulting and Clinical Psychology, Vol. 52, No. 6, 1984, p. 1039.

[69]Lund, J. P., Donga, R., Widmer, C. G., Stohler, C. S., "The pain-adaptation model: a discussion of the relationship between chronic musculoskeletal pain and motor activity", Canadian Journal of Physiology and Pharmacology, Vol. 69, No. 5, 1991, pp. 683–694.

[70]Svensson, P., Arendt-Nielsen, L., Houe, L., "Sensory-motor interactions of human experimental unilateral jaw muscle pain: a quantitative analysis", Pain, Vol. 64, No. 2, 1996, pp. 241–249.

[71]Falla, D., Farina, D., Dahl, M. K., Graven-Nielsen, T., "Muscle pain induces task-dependent changes in cervical agonist/antagonist activity", Journal of Applied Physiology, Vol. 102, No. 2, 2007, pp. 601–609.

[72]Farina, D., Arendt-Nielsen, L., Graven-Nielsen, T., "Experimental muscle pain decreases voluntary EMG activity but does not affect the muscle potential evoked by transcutaneous electrical stimulation", Clinical Neurophysiology, Vol. 116, No. 7, 2005, pp. 1558–1565.

[73]Graven-Nielsen, T., Svensson, P., Arendt-Nielsen, L., "Effects of experimental muscle pain on muscle activity and co-ordination during static and dynamic motor function", Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control, Vol. 105, No. 2, 1997, pp. 156–164.

[74]Magnusson, M., Aleksiev, A., Wilder, D., Pope, M., Spratt, K., Lee, S., Goel, V., Weinstein, J., "European spine society—the acromed prize for spinal research 1995 unexpected load and asymmetric posture as etiologic factors in low back pain", European Spine Journal, Vol. 5, No. 1, 1996, pp. 23–35.

[75]Radebold, A., Cholewicki, J., Panjabi, M. M., Patel, T. C., "Muscle response pattern to sudden trunk loading in healthy individuals and in patients with chronic low back pain", Spine, Vol. 25, No. 8, 2000, pp. 947–954.

[76]Tucker, K. J., Hodges, P. W., "Motoneurone recruitment is altered with pain induced in non-muscular tissue", Pain, Vol. 141, No. 1-2, 2009, pp. 151–155.

[77]Hodges, P., Richardson, C., "Feedforward contraction of transversus abdominis is not influenced by the direction of arm movement", Experimental Brain Research, Vol. 114, No. 2, 1997, pp. 362–370.

[78]Tsao, H., Galea, M., Hodges, P., "Reorganization of the motor cortex is associated with postural control deficits in recurrent low back pain", Brain, Vol. 131, No. 8, 2008, pp. 2161–2171.

[79]Fairbank, J., Couper, J., Davies, J., O'brien, J. *et al.*, "The oswestry low back pain disability questionnaire", Physiotherapy, Vol. 66, No. 8, 1980, pp. 271–273.

[80]Roland, M., Morris, R., "A study of the natural history of back pain. part i: Development of a reliable and sensitive measure of disability in low-back pain", Spine, Vol. 8, No. 2, 1983, pp. 141–144.

[81]Srhoj-Egekher, V., Cifrek, M., Medved, V., "The application of hilbert–huang transform in the analysis of muscle fatigue during cyclic dynamic contractions", Medical & Biological Engineering & Computing, Vol. 49, No. 6, 2011, pp. 659–669.

[82]Stango, A., Negro, F., Farina, D., "Spatial correlation of high density EMG signals provides features robust to electrode number and shift in pattern recognition for myocontrol", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 23, No. 2, 2015, pp. 189–198.

[83]Cifrek, M., Medved, V., Tonkovi ć, S., Ostojić, S., "Surface EMG based muscle fatigue evaluation in biomechanics", Clinical Biomechanics, Vol. 24, No. 4, 2009, pp. 327–340.

[84]Chowdhury, R. H., Reaz, M. B., Ali, M. A. B. M., Bakar, A. A., Chellappan, K., Chang, T. G., "Surface electromyography signal processing and classification techniques", Sensors, Vol. 13, No. 9, 2013, pp. 12 431–12 466.

[85]Phinyomark, A., Phukpattaranont, P., Limsakul, C., "Feature reduction and selection for EMG signal classification", Expert Systems with Applications, Vol. 39, No. 8, 2012, pp. 7420–7431.

[86]González-Izal, M., Malanda, A., Gorostiaga, E., Izquierdo, M., "Electromyographic models to assess muscle fatigue", Journal of Electromyography and Kinesiology, Vol. 22, No. 4, 2012, pp. 501–512.

[87]Hakonen, M., Piitulainen, H., Visala, A., "Current state of digital signal processing in myoelectric interfaces and related applications", Biomedical Signal Processing and Control, Vol. 18, 2015, pp. 334–359.

[88]Nazmi, N., Abdul Rahman, M., Yamamoto, S.-I., Ahmad, S., Zamzuri, H., Mazlan, S., "A review of classification techniques of EMG signals during isotonic and isometric contractions", Sensors, Vol. 16, No. 8, 2016, p. 1304.

[89]Ostoji ć, S., Srhoj-Egekher, V., Peharec, S., Cifrek, M., "A non-arbitrary definition of the analyzing interval of sEMG signal measured during sustained low back extension", in 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), May 2018, pp. 0245-0249.

[90]Peharec, S., Jerkovi ć, R., "Functional assessment in patients with non-specific low back pain", Medicina Fluminensis: Medicina Fluminensis, Vol. 50, No. 1, 2014, pp. 67–73.

[91]Merletti, R., De Luca, C. J., Sathyan, D., "Electrically evoked myoelectric signals in back muscles: effect of side dominance", Journal of Applied Physiology, Vol. 77, No. 5, 1994, pp. 2104–2114.

[92]Mannion, A. F., Connolly, B., Wood, K., Dolan, P., "The use of surface ENIG power spectral analysis in the evaluation of back muscle function", Development, Vol. 34, No. 4, 1997, pp. 427–439.

[93]Arendt-Nielsen, L., Graven-Nielsen, T., Svarrer, H., Svensson, P., "The influence of low back pain on muscle activity and coordination during gait: a clinical and experimental study", Pain, Vol. 64, No. 2, 1996, pp. 231–240.

[94]Ferguson, S. A., Marras, W. S., Burr, D. L., Davis, K. G., Gupta, P., "Differences in motor recruitment and resulting kinematics between low back pain patients and asymptomatic

participants during lifting exertions", Clinical Biomechanics, Vol. 19, No. 10, 2004, pp. 992–999.

[95]Takemasa, R., Yamamoto, H., Tani, T., "Trunk muscle strength in and effect of trunk muscle exercises for patients with chronic low back pain. the differences in patients with and without organic lumbar lesions.", Spine, Vol. 20, No. 23, 1995, pp. 2522–2530.

[96]Lee, J.-H., Ooi, Y., Nakamura, K., "Measurement of muscle strength of the trunk and the lower extremities in subjects with history of low back pain.", Spine, Vol. 20, No. 18, 1995, pp. 1994–1996.

[97]Wong, K. W., Leong, J. C., Chan, M.-k., Luk, K. D., Lu, W. W., "The flexion–extension profile of lumbar spine in 100 healthy volunteers", Spine, Vol. 29, No. 15, 2004, pp. 1636–1641.

[98]Hoyt, W., Hunt Jr, H., De Pauw, M., Bard, D., Shaffer, F., Passias, J., Robbins Jr, D., Runyon, D., Semrad, S., Symonds, J. *et al.*, "Electromyographic assessment of chronic low-back pain syndrome.", The Journal of the American Osteopathic Association, Vol. 80, No. 11, 1981, p. 728.

[99]Roy, S. H., De Luca, C. J., Snyder-Mackler, L., Emley, M. S., Crenshaw, R. L., Lyons, J. P., "Fatigue, recovery, and low back pain in varsity rowers", Med Sci Sports Exerc, Vol. 22, No. 4, 1990, pp. 463–9.

[100]Farina, D., Merletti, R., "Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions", Journal of Electromyography and Kinesiology, Vol. 10, No. 5, 2000, pp. 337–349.

[101]Karlsson, S., Yu, J., Akay, M., "Time-frequency analysis of myoelectric signals during dynamic contractions: a comparative study", IEEE Transactions on Biomedical Engineering, Vol. 47, No. 2, 2000, pp. 228–238.

[102]Knaflitz, M., Bonato, P., "Time-frequency methods applied to muscle fatigue assessment during dynamic contractions", Journal of Electromyography and Kinesiology, Vol. 9, No. 5, 1999, pp. 337–350.

[103]Oppenheim, A. V., Discrete-Time Signal Processing. Pearson Education India, 1999.

[104]Farina, D., "Interpretation of the surface electromyogram in dynamic contractions", Exercise and Sport Sciences Reviews, Vol. 34, No. 3, 2006, pp. 121–127.

[105]Farina, D., Fattorini, L., Felici, F., Filligoi, G., “Nonlinear surface EMG analysis to detect changes of motor unit conduction velocity and synchronization”, Journal of Applied Physiology, Vol. 93, No. 5, 2002, pp. 1753–1763.

[106]Lei, M., Wang, Z., Feng, Z., “Detecting nonlinearity of action surface EMG signal”, Physics Letters A, Vol. 290, No. 5-6, 2001, pp. 297–303.

[107]Nieminen, H., Takala, E., “Evidence of deterministic chaos in the myoelectric signal”, Electromyography and Clinical Neurophysiology, Vol. 36, 1996, pp. 49–58.

[108]MacIsaac, D. T., Parker, P. A., Englehart, K. B., Rogers, D. R., “Fatigue estimation with a multivariable myoelectric mapping function”, IEEE Transactions on Biomedical Engineering, Vol. 53, No. 4, 2006, pp. 694–700.

[109]Chua, K. C., Chandran, V., Acharya, U. R., Lim, C. M., “Application of higher order statistics/spectra in biomedical signals—a review”, Medical Engineering & Physics, Vol. 32, No. 7, 2010, pp. 679–689.

[110]Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., Liu, H. H., “The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis”, in Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, Vol. 454, No. 1971. The Royal Society, 1998, pp. 903–995.

[111]Huang, N. E., Hilbert-Huang Transform and Its Applications. World Scientific, 2014, Vol. 16.

[112]Dragomiretskiy, K., Zosso, D., “Variational mode decomposition”, IEEE Transactions on Signal Processing, Vol. 62, No. 3, 2014, pp. 531–544.

[113]Colominas, M. A., Schlotthauer, G., Torres, M. E., “Improved complete ensemble emd: A suitable tool for biomedical signal processing”, Biomedical Signal Processing and Control, Vol. 14, 2014, pp. 19–29.

[114]Andrade, A. O., Nasuto, S., Kyberd, P., Sweeney-Reed, C. M., Van Kanijn, F., “EMG signal filtering based on empirical mode decomposition”, Biomedical Signal Processing and Control, Vol. 1, No. 1, 2006, pp. 44–55.

[115]Jovi ć, A., Stančin, I., Friganović, K., Cifrek, M., “Clinical decision support systems in practice: Current status and challenges”, in 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2020, pp. 355–360.

[116]Devillé, W. L., van der Windt, D. A., Dzaferagic, A., Bezemer, P., Bouter, L. M., "The test of lasegue: systematic review of the accuracy in diagnosing herniated discs", Spine, Vol. 25, No. 9, 2000, pp. 1140–1147.

[117]Willhuber, G. O. C., Piuzzi, N. S., "Straight leg raise test", in StatPearls [Internet]. StatPearls Publishing, 2021.

[118]Hermens, H. J., Freriks, B., Merletti, R., Stegeman, D., Blok, J., Rau, G., Disselhorst-Klug, C., Hägg, G., "European recommendations for surface electromyography", Roessingh research and development, Vol. 8, No. 2, 1999, pp. 13–54.

[119]Merletti, R., Di Torino, P., "Standards for reporting EMG data", Journal of Electromyography and Kinesiology, Vol. 9, No. 1, 1999, pp. 3–4.

[120]Merletti, R., Conte, L. L., "Advances in processing of surface myoelectric signals: Part 1", Medical and Biological Engineering and Computing, Vol. 33, No. 3, 1995, pp. 362–372.

[121]Beck, T. W., Housh, T. J., Johnson, G. O., Weir, J. P., Cramer, J. T., Coburn, J. W., Malek, M. H., "Comparison of fourier and wavelet transform procedures for examining the mechanomyographic and electromyographic frequency domain responses during fatiguing isokinetic muscle actions of the biceps brachii", Journal of Electromyography and Kinesiology, Vol. 15, No. 2, 2005, pp. 190–199.

[122]MacIsaac, D., Parker, P., Scott, R., Englehart, K., Duffley, C., "Influences of dynamic factors on myoelectric parameters.", IEEE engineering in medicine and biology magazine: the quarterly magazine of the Engineering in Medicine & Biology Society, Vol. 20, No. 6, 2001, pp. 82–89.

[123]Hudgins, B., Parker, P., Scott, R. N., "A new strategy for multifunction myoelectric control", IEEE Transactions on Biomedical Engineering, Vol. 40, No. 1, 1993, pp. 82–94.

[124]Du, Y.-C., Lin, C.-H., Shyu, L.-Y., Chen, T., "Portable hand motion classifier for multi-channel surface electromyography recognition using grey relational analysis", Expert Systems with Applications, Vol. 37, No. 6, 2010, pp. 4283–4291.

[125]Olofsen, E., Sleigh, J., Dahan, A., "Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect", British Journal of Anaesthesia, Vol. 101, No. 6, 2008, pp. 810–821.

[126]Xi, X., Tang, M., Miran, S. M., Luo, Z., "Evaluation of feature extraction and recognition for activity monitoring and fall detection based on wearable sEMG sensors", Sensors, Vol. 17, No. 6, 2017, p. 1229.

[127]Bandt, C., Pompe, B., "Permutation entropy: a natural complexity measure for time series", Physical Review Letters, Vol. 88, No. 17, 2002, p. 174102.

[128]Unakafova, V. A., Keller, K., "Efficiently measuring complexity on the basis of real-world data", Entropy, Vol. 15, No. 10, 2013, pp. 4392–4415.

[129]Aggarwal, C. C., Hinneburg, A., Keim, D. A., "On the surprising behavior of distance metrics in high dimensional space", in International Conference on Database Theory. Springer, 2001, pp. 420–434.

[130]Sakoe, H., Chiba, S., "Dynamic programming algorithm optimization for spoken word recognition", IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 26, No. 1, 1978, pp. 43–49.

[131]Ratanamahatana, C. A., Keogh, E., "Making time-series classification more accurate using learned constraints", in Proceedings of the 2004 SIAM International Conference on Data Mining. SIAM, 2004, pp. 11–22.

[132]Dhingra, S. D., Nijhawan, G., Pandit, P., "Isolated speech recognition using MFCC and DTW", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2, No. 8, 2013, pp. 4085–4092.

[133]Kale, A., Cuntoor, N., Yegnanarayana, B., Rajagopalan, A., Chellappa, R., "Gait analysis for human identification", in International Conference on Audio-and Video-Based Biometric Person Authentication. Springer, 2003, pp. 706–714.

[134]Serra, J., Arcos, J. L., "An empirical evaluation of similarity measures for time series classification", Knowledge-Based Systems, Vol. 67, 2014, pp. 305–314.

[135]Keogh, E., "Introduction to Time Series Mining", available on: http://didawiki.cli.di.unipi.it/lib/exe/fetch.php/dm/time_series_from_keogh_tutorial.pdf 2019.

[136]Myers, J. L., Well, A., Lorch, R. F., Research Design and Statistical Analysis, 3rd ed. New York, NY, USA: Routledge, 2010.

[137]Daniel, W. W. *et al.*, Applied Nonparametric Statistics. Boston, MA, USA: Houghton Mifflin, 1978.

[138]Al-Angari, H. M., Kanitz, G., Tarantino, S., Cipriani, C., "Distance and mutual information methods for EMG feature and channel subset selection for classification of hand movements", Biomedical Signal Processing and Control, Vol. 27, 2016, pp. 24–31.

[139]Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Philip, S. Y. *et al.*, "Top 10 algorithms in data mining", Knowledge and Information Systems, Vol. 14, No. 1, 2008, pp. 1–37.

[140]Gibbons, J. D., Chakraborti, S., Nonparametric Statistical Inference, 6th ed. New York, NY, USA: Chapman and Hall/CRC press, 2020.

[141]Goldberger, J., Hinton, G. E., Roweis, S. T., Salakhutdinov, R. R., "Neighbourhood components analysis", in Advances in Neural Information Processing Systems, 2005, pp. 513–520.

[142]Yang, W., Wang, K., Zuo, W., "Neighborhood component feature selection for high-dimensional data.", Journal of Computers, Vol. 7, No. 1, 2012, pp. 161–168.

[143]Weinberger, K. Q., Saul, L. K., "Distance metric learning for large margin nearest neighbor classification.", Journal of Machine Learning Research, Vol. 10, No. 2, 2009.

[144]MATLAB version 9.9.0.1524771 (R2020b) Update 2, The Mathworks, Inc., Natick, Massachusetts, 2020.

[145]Breiman, L., Friedman, J., Olshen, R., Stone, C., Classification and Regression Trees. New York, NY, USA: Routledge, 1984.

[146]Fisher, R. A., "The use of multiple measurements in taxonomic problems", Annals of Eugenics, Vol. 7, No. 2, 1936, pp. 179–188.

[147]Fukunaga, K., Introduction to Statistical Pattern Recognition, 2nd ed. San Diego, CA, USA: Academic Press, 1990.

[148]Dreiseitl, S., Ohno-Machado, L., "Logistic regression and artificial neural network classification models: a methodology review", Journal of Biomedical Informatics, Vol. 35, No. 5-6, 2002, pp. 352–359.

[149]Cristianini, N., Shawe-Taylor, J. *et al.*, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge, UK: Cambridge University Press, 2000.

[150]Vapnik, V. N., The Nature of Statistical Learning Theory, 2nd ed. New York, NY, USA: Springer, 2013.

[151]Fix, E., Hodges, J. L., "Discriminatory analysis. nonparametric discrimination: Consistency properties", International Statistical Review/Revue Internationale de Statistique, Vol. 57, No. 3, 1989, pp. 238–247.

[152]Altman, N. S., “An introduction to kernel and nearest-neighbor nonparametric regression”, The American Statistician, Vol. 46, No. 3, 1992, pp. 175–185.

[153]Breiman, L., “Bagging predictors”, Machine Learning, Vol. 24, No. 2, 1996, pp. 123–140.

[154]Schapire, R. E., “The strength of weak learnability”, Machine Learning, Vol. 5, No. 2, 1990, pp. 197–227.

[155]Freund, Y., Schapire, R. E., “A decision-theoretic generalization of on-line learning and an application to boosting”, Journal of Computer and System Sciences, Vol. 55, No. 1, 1997, pp. 119–139.

[156]Friedman, J., Hastie, T., Tibshirani, R. *et al.*, “Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors)”, The Annals of Statistics, Vol. 28, No. 2, 2000, pp. 337–407.

[157]Ho, T. K., “The random subspace method for constructing decision forests”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 8, 1998, pp. 832–844.

[158]Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., Napolitano, A., “Rusboost: Improving classification performance when training data is skewed”, in 2008 19th International Conference on Pattern Recognition. IEEE, 2008, pp. 1–4.

[159]Hastie, T., Tibshirani, R., Friedman, J., The Elements of Statistical Learning, 2nd ed. Springer, New York, NY.

[160]Tharwat, A., “Linear vs. quadratic discriminant analysis classifier: a tutorial”, International Journal of Applied Pattern Recognition, Vol. 3, No. 2, 2016, pp. 145–180.

[161]Ng, A. Y., Jordan, M. I., “On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes”, in Advances in Neural Information Processing Systems, 2002, pp. 841–848.

[162]Greene, W. H., “Econometric analysis”, Journal of Boston: Pearson Education, 2012, pp. 803–806.

[163]Guler, I., Ubeyli, E. D., “Multiclass support vector machines for eeg-signals classification”, IEEE Transactions on Information Technology in Biomedicine, Vol. 11, No. 2, 2007, pp. 117–126.

[164]García-Gonzalo, E., Fernández-Muñiz, Z., García Nieto, P. J., Bernardo Sánchez, A., Menéndez Fernández, M., “Hard-rock stability analysis for span design in entry-type excavations with learning classifiers”, Materials, Vol. 9, No. 7, 2016, p. 531.

[165]Han, E.-H. S., Karypis, G., Kumar, V., “Text categorization using weight adjusted k-nearest neighbor classification”, in Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, 2001, pp. 53–65.

[166]Breiman, L., “Random forests”, Machine Learning, Vol. 45, No. 1, 2001, pp. 5–32.

[167]Ambroise, C., McLachlan, G. J., “Selection bias in gene extraction on the basis of microarray gene-expression data”, Proc. Natl. Acad. Sci. U.S.A., Vol. 99, No. 10, 2002, pp. 6562-6566.

[168]Smialowski, P., Frishman, D., Kramer, S., “Pitfalls of supervised feature selection”, Bioinformatics, Vol. 26, No. 3, 2010, pp. 440–443.

[169]Mannion, A. F., Weber, B. R., Dvorak, J., Grob, D., Müntener, M., “Fibre type characteristics of the lumbar paraspinal muscles in normal healthy subjects and in patients with low back pain”, Journal of Orthopaedic Research, Vol. 15, No. 6, 1997, pp. 881–887.

[170]Martinez, W. L., Martinez, A. R., Solka, J., Exploratory Data Analysis with MATLAB. Chapman and Hall/CRC, 2017.

[171]Fu, T.-c., “A review on time series data mining”, Engineering Applications of Artificial Intelligence, Vol. 24, No. 1, 2011, pp. 164–181.

[172]Riveros, N. A. M., Espitia, B. A. C., Pico, L. E. A., “Comparison between k-means and self-organizing maps algorithms used for diagnosis spinal column patients”, Informatics in Medicine Unlocked, Vol. 16, 2019, p. 100206.

[173]Maimon, O., Rokach, L., “Introduction to knowledge discovery and data mining”, in Data Mining and Knowledge Discovery Handbook. Springer, 2009, pp. 1–15.

[174]Ward Jr, J. H., “Hierarchical grouping to optimize an objective function”, Journal of the American Statistical Association, Vol. 58, No. 301, 1963, pp. 236–244.

[175]Zahn, C. T., “Graph-theoretical methods for detecting and describing gestalt clusters”, IEEE Transactions on Computers, Vol. 100, No. 1, 1971, pp. 68–86.

[176]Hubley-Kozey, C., Vezina, M., “Differentiating temporal electromyographic waveforms between those with chronic low back pain and healthy controls”, Clinical Biomechanics, Vol. 17, No. 9-10, 2002, pp. 621–629.

[177]Oddsson, L. I., De Luca, C. J., “Activation imbalances in lumbar spine muscles in the presence of chronic low back pain”, Journal of Applied Physiology, Vol. 94, No. 4, 2003, pp. 1410–1420.

[178]O'Sullivan, P., "Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism", Manual Therapy, Vol. 10, No. 4, 2005, pp. 242–255.

[179]Van Dieën, J. H., Reeves, N. P., Kawchuk, G., Van Dillen, L. R., Hodges, P. W., "Analysis of motor control in patients with low back pain: a key to personalized care?", Journal of Orthopaedic & Sports Physical Therapy, Vol. 49, No. 6, 2019, pp. 380–388.

[180]Lu, W. W., Luk, K. D., Cheung, K. M., Wong, Y. W., Leong, J. C., "Back muscle contraction patterns of patients with low back pain before and after rehabilitation treatment: an electromyographic evaluation", Clinical Spine Surgery, Vol. 14, No. 4, 2001, pp. 277–282.

[181]Reeves, N. P., Cholewicki, J., Silfies, S. P., "Muscle activation imbalance and low-back injury in varsity athletes", Journal of Electromyography and Kinesiology, Vol. 16, No. 3, 2006, pp. 264–272.

Biography

Vedran Srhoj-Egekher was born in Dubrovnik, Croatia, in 1984. During his pre-university education, he took part in many sports (chess, football, handball) and knowledge competitions (math, physics, chemistry) on the local, county, and national levels. As a reward for his numerous achievements during high school, he was awarded direct enrollment in the Faculty of Electrical Engineering and Computing (FER) at the University of Zagreb (2003). He received his Dipl.-Ing. degree in electrical engineering from the University of Zagreb, in 2009 as one of the best-in-generation students in the electrical engineering program (within 3% of best students). His graduate thesis "The analysis of surface myoelectric signals recorded during dynamic contraction" was awarded as the best graduate thesis, for the academic year 2008/2009, in the interdisciplinary field of biomedical engineering by Croatian Medical and Biological Engineering Society (CROMBES) in 2010. Directly after graduating, end of 2009, he starts his job as a software developer at the private ICT company Ericsson Nikola Tesla d.d. (Zagreb). During his thirteen years career in Ericson Nikola Tesla, he takes different positions and responsibilities in software development and solutions for advanced mobile networks rollout tools. In 2017 he takes the position of manager, with the responsibility of running a global software development unit. In 2017, he enrolled post-graduate program at the University of Zagreb, pursuing his Ph.D. degree in the field of electrical engineering with interest on biomedical signal analysis and machine learning applications in medicine. He is an IEEE member, the first author and co-author of several conference and scientific papers, some also produced in a collaboration with one of the top-ranking international scientific groups for medical image analysis, Image Sciences Institute, University Medical Center Utrecht. Vedran is the first author or co-author of scientific papers with $\sim$100 citations according to WoS (92), Scopus (112) and Google Scholar (176) databases. He is also a member of several associations within the Croatian startups ecosystem and is currently acting as one of the mentors at the international EU DigiTour project to build the capacities and competencies of experts in tourism in the domain of artificial intelligence. He also holds several certificates in business consulting and software solutioning.

List of published works

Journal publications

1.Srhoj-Egekher, V., Cifrek, M., Peharec, S., "Feature Modeling for Interpretable Low Back Pain Classification Based on Surface EMG", IEEE Access Vol. 10, 2022, pp. 73702–73727.

2.Ostoji ć, S., Peharec, S., Srhoj-Egekher, V., Cifrek, M., "Differentiating patients with radiculopathy from chronic low back pain patients by single surface EMG parameter", Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije. Vol. 59, No. 3-4, 2018, pp. 400–407.

3.Išgum, Ivana, et al., "Evaluation of automatic neonatal brain segmentation algorithms: the NeoBrainS12 challenge", Medical Image Analysis, Vol. 20, No. 1, 2015, pp. 135–151.

4.Srhoj-Egekher, V., Cifrek, M., Medved, V., "The application of Hilbert–Huang transform in the analysis of muscle fatigue during cyclic dynamic contractions". Medical & Biological Engineering & Computing, Vol. 49, No. 6, 2011, pp. 659–669.

Conference papers

1.Srhoj-Egekher, V., Cifrek, M., Peharec, S., "Low-back pain patients classification based on sEMG activation patterns detection", in 1st European Congress on Biomedical and Veterinary Engineering (BioMedVetMech), October 2022.

2.Ostoji ć, S., Srhoj-Egekher, V., Peharec, S., Cifrek, M., "A non-arbitrary definition of the analyzing interval of sEMG signal measured during sustained low back extension", in 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, May 2018, pp. 0245–0249.

3.Srhoj-Egekher, V., Benders, M. J. N. L., Viergever, M. A., Išgum, I., "Automatic neonatal brain tissue segmentation with MRI", in Proc. SPIE. Vol. 8669, 2013, pp. 426–433.

4.Srhoj-Egekher, V., Benders, M. J. N. L., Kersbergen, K. J., Viergever, M. A., Isgum, I., "Automatic segmentation of neonatal brain MRI using atlas based segmentation and machine learning approach", in MICCAI Grand Challenge on Neonatal Brain Segmentation 2012 (NeoBrainS12), 2012, pp. 22-–27.

5.Srhoj-Egekher V., Peharec S., Cifrek M., "Surface EMG frequency-based discrimination of low back pain patients", in WC2012 World Congress on Medical Physics and Biomedical Engineering Proceedings, Beijing, Editor: Liu D., 2012.

Životopis

Vedran Srhoj-Egekher je rođen u Dubrovniku, Hrvatska, 1984. godine. Tijekom svog osnovno i srednjoškolskog obrazovanja sudjeluje na brojnim športskim natjecanjima (šah, nogomet, rukomet) i natjecanjima u znanju (matematika, fizika, kemija) na lokalnoj, županijskoj i državnoj razini, te osvaja brojne nagrade i pohvale. Kao nagrada za navedene uspjehe, Vedran stječe pravo na upis na Fakultet elektrotehnike i računarstva (FER), Sveučilište u Zagrebu, bez prijamnog ispita, na koji se i upisuje 2003. godine. Svoju titulu dipl. ing. elektrotehnike je stekao diplomiravši 2009. godine kao jedan od najboljih studenata u generaciji (unutar 3% najboljih studenata). Njegov diplomski rad "Analiza površinskih mioelektričnih signala izmjerenih tijekom dinamičkih kontrakcija" nagrađen je kao najbolji diplomski rad, za akademsku godinu 2008./2009., u interdisciplinarnom području biomedicinskog inženjerstva od strane Hrvatskog društva za medicinsku i biološku tehniku (CROMBES) 2010. godine. Po završetku studija, krajem 2009. god., zapošljava se kao softver developer u privatnoj ICT kompaniji Ericsson Nikola Tesla d.d. (Zagreb). Tijekom svoje trinaestogodišnje karijere u tvrtki Ericson Nikola Tesla, preuzima različite pozicije i odgovornosti u razvoju softvera i rješenja za napredne alate za uvođenje mobilnih mreža. Krajem 2017. god. preuzima poziciju menadžera odjela za razvoj softvera u sklopu globalnih kompanijskih odgovornosti. Godine 2017. upisuje poslijediplomski doktorkski studij na Sveučilištu u Zagrebu, područje elektrotehnike, s interesom za analizu biomedicinskih signala i primjene strojnog učenja u medicini. Član je IEEE, prvi je autor i suautor nekoliko konferencijskih i znanstvenih radova, od kojih neki također nastali u suradnji s jednom od vodećih međunarodnih znanstvenih grupa za analizu medicinskih slika, Image Sciences Institute, University Medical Center Utrecht. Vedran je prvi autor ili suautor znanstvenih radova s $\sim$100 citata prema bazama WoS (92), Scopus (112) i Google Scholar (176). Također je član nekoliko udruga unutar hrvatskog startup ekosustava, a izabran je i za jednog od mentora na međunarodnom projektu EU DigiTour za izgradnju kapaciteta i kompetencija stručnjaka u turizmu u domeni umjetne inteligencije. Također posjeduje nekoliko certifikata za poslovno savjetovanje i projektiranje softverskih rješenja.