ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

TSALLIS ENTROPY BASED FEATURE EXTRACTION FROM INSOLE FORCE SENSOR DATA TO DIAGNOSE VESTIBULAR SYSTEM DISORDERS

M.Sc. THESIS

Harun Yaşar KÖSE

Department of Mechatronics Engineering

Mechatronics Engineering Programme

DECEMBER 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

TSALLIS ENTROPY BASED FEATURE EXTRACTION FROM INSOLE FORCE SENSOR DATA TO DIAGNOSE VESTIBULAR SYSTEM DISORDERS

M.Sc. THESIS

Harun Yaşar KÖSE
(518211015)

Department of Mechatronics Engineering

Mechatronics Engineering Programme

Thesis Advisor: Assoc. Prof. Dr. Serhat İKİZOĞLU

DECEMBER 2023

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

VESTİBÜLER SİSTEM BOZUKLUKLARININ TANISI İÇİN TABANLIK KUVVET ALGILAYICILARI VERİLERİNDEN TSALLİS ENTROPİSİ TABANLI ÖZNİTELİK ÇIKARIMI

YÜKSEK LİSANS TEZİ

Harun Yaşar KÖSE
(518211015)

Mekatronik Mühendisliği Anabilim Dalı

Mekatronik Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Serhat İKİZOĞLU

ARALIK 2023

Harun Yaşar KÖSE, a M.Sc. student of İTU Graduate School, student ID 518211015, successfully defended the thesis/dissertation entitled "TSALLIS ENTROPY BASED FEATURE EXTRACTION FROM INSOLE FORCE SENSOR DATA TO DIAGNOSE VESTIBULAR SYSTEM DISORDERS", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assoc. Prof. Dr. Serhat İKİZOĞLU** İstanbul Technical University	....
Jury Members :	**Prof. Dr. Fikret ÇALIŞKAN** İstanbul Technical University	....
	Prof. Dr. Süleyman Hakan ÜNDİL İstanbul Nişantaşı University	....

Date of Submission : 15 December 2023
Date of Defense : 21 December 2023

To my lovely family,

FOREWORD

I am deeply grateful to Assoc. Prof. Dr. Serhat İKİZOĞLU, my esteemed thesis advisor, for his invaluable guidance and unwavering support throughout the research process. His expertise and encouragement were pivotal in shaping the trajectory of this study.

A separate heartfelt appreciation goes to my mother and sister, Nadime and Elif, for their unwavering support throughout this academic journey. Their love and encouragement provided the foundation upon which this research was built.

In acknowledging the collaborative efforts that have contributed to the realization of this work, I express sincere gratitude to those involved in data collection and the medical personnel who facilitated the project. Special thanks are also extended to Prof. Dr. Ahmet Ataş, Assoc. Prof. Dr. Eyüp Kara, Saddam Heyderov, and Tunay Çakar for their invaluable contributions.

I consider myself truly fortunate to have been surrounded by such dedicated and supportive individuals, each playing a significant role in the successful completion of this research.

December 2023

Harun Yaşar KÖSE
Mechatronics Eng.

TABLE OF CONTENTS

Page

ABBREVIATIONS

ALS	: Amyotrophic Lateral Sclerosis
BPPV	: Benign Paroxysmal Positional Vertigo
CPD	: Computerized Dynamic Posturography
DT	: Decision Trees
HD	: Huntington's Disease
KNN	: Nearest Neighbors
MAE	: Multiscale Approximate Entropy
MATLAB	: Matrix Laboratory
PD	: Parkinson's Diseases
ROC	: Receiver Operating Characteristic
SE	: Shannon Entropy
SVM	: Support Vector Machine
TE	: Tsallis Entropy
VS	: Vestibular System

SYMBOLS

$\boldsymbol{S_q}$	: Tsallis entropy value
$\boldsymbol{q}$	: Generalization parameter of Tsallis entropy
$\boldsymbol{W}$	: Number of possible states for probabilities
$\boldsymbol{p_i}$	: Probability of the i-th state
$\boldsymbol{w}$	: Relationship between the weight/force [N]
$\boldsymbol{v_o}$	: Output voltage [V]
$\boldsymbol{S_{Rmax}}, \boldsymbol{S_{Lmax}}$	: An ordered set containing the maximum values of time series data of sensors. (R: Right, L: Left)
$\boldsymbol{S_k}$	: The time series data of the sensor no x on the insoles (For k = 0, 1, 2, …, 7)
$\boldsymbol{F_i}$	: Time series data of i-th order step data
$\boldsymbol{T_i}$	: Trend of time series data of i-th order step data
$\boldsymbol{\breve{T}_i}$	: Trend of time series data of i-th order step data whose data length is scaled
$\boldsymbol{i}$	: Step sequence in the time series data
$\boldsymbol{n}$	: Total number of steps
$\boldsymbol{\alpha}$	: Weighted average coefficient
$\boldsymbol{E}$	: Ordered set of step-to-step entropy values
$\boldsymbol{E'}$	: Augmented E with negative counterparts
$\boldsymbol{\sigma(E')}$	: Standard deviation of set E'

LIST OF TABLES

Page

LIST OF FIGURES

Page

TSALLIS ENTROPY BASED FEATURE EXTRACTION FROM INSOLE FORCE SENSOR DATA TO DIAGNOSE VESTIBULAR SYSTEM DISORDERS

SUMMARY

The vestibular system plays a crucial role in maintaining an individual's ability to carry out daily activities independently and safely. This study proposes a novel approach using Tsallis entropy analysis of insole force sensor data to identify diseases associated with vestibular system dysfunction. By analyzing Tsallis entropy values of the entire gait cycle and change of Tsallis entropy from step-to-step, the study aims to differentiate between individuals with good health and those with vestibular system-related diseases.

A notable finding of the study is the observation that the histogram of normalized and interpolated sensor data contains fewer bins for healthy subjects. This reduction in the number of bins can be attributed to improved balance and coordination, which leads to reduced fluctuation around the trend curve. Unlike previous studies that focus on gait dynamics and require extensive walking time, this research takes a different approach by directly processing instantaneous force values to extract features. One key innovation of this research is the development of a specifically designed algorithm to generate the trend curve. This algorithm enables the extraction of significant insights even from relatively short walking sessions. By applying this algorithm, the study successfully extracts a feature set from the force sensor data.

The extracted feature set is then inputted into fundamental classification algorithms. Among these algorithms, the Support Vector Machine (SVM) demonstrates the highest performance. It achieves an average accuracy of 95% in binary classification, effectively distinguishing between healthy individuals and those suffering from vestibular system-related diseases. This high accuracy indicates the potential of Tsallis entropy analysis as a valuable tool in identifying diseases associated with vestibular system dysfunction.

This study represents a significant milestone within a comprehensive project aimed at identifying distinct vertigo syndromes associated with balance disorders and determining their respective stages, if applicable. The exceptional performance observed in this study serves as a compelling impetus for further exploration and inquiry into this matter.

In conclusion, this study introduces a novel approach utilizing Tsallis entropy analysis of insole force sensor data to identify diseases related to vestibular system dysfunction. By extracting features and employing classification algorithms, the study successfully distinguishes between healthy individuals and those with vestibular system-related diseases with a high accuracy of 95%. This research presents a significant contribution to the broader project of identifying vertigo syndromes and determining their stage, showcasing the potential of Tsallis entropy analysis in understanding and addressing balance disorders associated with the vestibular system.

VESTİBÜLER SİSTEM BOZUKLUKLARININ TANISI İÇİN TABANLIK KUVVET ALGILAYICILARI VERİLERİNDEN TSALLİS ENTROPİSİ TABANLI ÖZNİTELİK ÇIKARIMI

ÖZET

Vestibüler sistem, günlük yaşantımızda denge ve mekansal yönelimi sürdürmede anahtar bir role sahiptir. İç kulakta yer alan bu kritik yapı, hareketlerimizi ve etrafımızdaki dünyayı nasıl algıladığımızı koordine eder. Ancak, bu sistemdeki herhangi bir bozukluk, denge ve mekansal oryantasyon yeteneklerimizi ciddi şekilde etkileyebilir. Bu tip bozukluklar, bireyin günlük hayattaki aktivitelerini ve sosyal etkileşimlerini kısıtlayabilir. Vestibüler bozuklukların hızla ve doğru bir şekilde tanılanıp etkin bir şekilde izlenmesinin kritik önemi bu bağlamda daha da belirginleşir. Bu tez, vestibüler sistem bozukluklarından kaynaklanabilecek denge sorunlarının tespit edilmesine odaklanmıştır.

Araştırmada, vestibüler sistem bozuklukları olan bireylerde yürüyüşün nasıl etkilendiğini anlamak için yürüyüş sırasında ayak tabanına uygulanan kuvvetin analizi ele alınmıştır. Bu bağlamda, kuvvet verilerini detaylıca analiz etmek için Tsallis entropisi kullanılmıştır. Tsallis entropisinin yürüyüş dinamiklerinin karmaşıklığını aydınlatmada büyük bir potansiyele sahip olduğu gözlemlenmiştir. Uygulanan bu yaklaşımın, yürüyüş örüntülerindeki vestibüler bozukluklardan kaynaklanan belirgin değişiklikleri tanımlamada önemli bir araç olduğu bulgusuna erişilmiştir.

Entropi, dinamik bir sistemin içsel düzensizliğini ve rastgeleliğini kantitatif olarak değerlendirmek için temel bir parametre olarak kabul edilmektedir. Shannon entropisi, kısa vadeli mikroskobik korelasyonlara sahip sistemlerin yapısını tanımlamada etkilidir. Tsallis entropisi, Shannon entropisinin genelleştirilmiş bir varyasyonu olarak, sistem karmaşıklığı ve yapısal özelliklerin kantitatif analizinde önemli bir role sahiptir. Uzun süreli etkileşim karakteristiği gösteren sistemlerde, Tsallis entropisinin genelleştirilmiş yaklaşımı, zaman serilerinin içerdiği gizli bilgilere detaylı bir şekilde erişebilir; bu, geleneksel yöntemlerle elde edilemeyen önemli içgörüler sağlar.

Vestibüler sistemle ilişkili yürüyüş verisinin analizi, karmaşık biyolojik sistemlerin doğası gereği doğru bilgi çıkarımı için karmaşık analiz yöntemlerine ihtiyaç duyar. Yürüyüş parametrelerinin analizinde Tsallis entropisinin kullanılması, bu yöntemin biyolojik sistemlerdeki doğrusal olmayan ve uzun menzilli korelasyonları ele alabilme yeteneğinden ileri gelmektedir. Tsallis entropisi, bu kompleks etkileşimleri esnek bir şekilde karakterize edebilir. Bir örnek olarak, literatürde elektroensefalografi (EEG) sinyallerinin analizinde Tsallis entropisi yöntemi kullanılmaktadır. Bu örnekte Tsallis entropisi, anormal beyin aktivite modellerini tanımlamada ve farklı beyin durumları arasında ayrım yapmada olumlu sonuçlar sağlamaktadır. Bu bağlamda, Tsallis entropisi yöntemi, yürüyüş verisi analizi gibi karmaşık sistemlerde bilgi içeriğini değerlendirme imkânı sağlar; bu da araştırmacılara yürüyüşün özelliklerine ve farklı durumlardaki bireylerdeki varyasyonlarına dair değerli bilgiler kazandırır.

Literatürde yer alan araştırmalara göre, yürüyüş verisinin analizlerinde, ayak tabanlarındaki ağırlık dağılımının dört ana noktada belirgin bir etkisi olduğu tespit edilmiştir. Bu dört ana nokta göz önünde bulundurularak, tabanlıklarda kuvvet algılayıcıları stratejik bir şekilde yerleştirilmiştir. Yürüyüş sırasında her bireyden, her ayakta dört, toplamda ise sekiz algılayıcı kullanılarak veri alınmıştır. Bu veri, tabanlık kuvvet algılayıları aracılığıyla bir Arduino Mega ünitesine kaydedilmiş ve ardından bir HC-06 Bluetooth modülü ile bir diz üstü bilgisayara aktarılmıştır. Ölçüm esnasında, her algılayıcıdan saniyede 20 örnek alınarak veri toplama gerçekleştirilmiştir. Tüm bu süreçte, odyologların görüş ve önerileri doğrultusunda çalışmalar yürütülmüştür. Katılan bireyler hakkında ağırlık, yaş gibi çalışmamıza yönelik detaylı bilgiler alınarak, hastalıklarının dağılımı dikkatlice kaydedilip değerlendirilmiştir. Araştırma sürecinde katılımcıların gizliliği titizlikle korunmuş ve etik ilkelere sıkıca riayet edilmiştir.

Özellikle yürüyüş verilerinin değerlendirilmesi gibi karmaşık analizlerde, veri işlemenin rolü, detaylı içgörü kazandırmak açısından kritiktir. Bu bağlamda, bu çalışmanın kapsamı dahilinde uygulanan veri işleme süreci, optimize edilmiş altı aşamalı bir süreç olarak tanımlanmıştır. İlk aşama olan normalize etme aşamasında, ham algılayıcı verileri, farklı bireyler arasında karşılaştırılabilir olmaları amacıyla normalize edilmiştir. Bu süreçte, analize uygun olmayan kısımlar veri setlerinden çıkarılmıştır. İkinci aşamada adım analizi gerçekleştirilmiştir. Bu aşamada, bireyin sağ ve sol ayağı için algılayıcılardan elde edilen veriler iki grup halinde değerlendirilmiştir. Geliştirilen algoritmalar yardımıyla, yürüme sırasında ayağın zeminle temas ettiği ve zeminden ayrıldığı anlar tespit edilmiştir. Daha sonrasında sensör verileri, yürüyüş adımlarına göre segmentlere ayrılmış, bu da yürüyüş modelinin daha ayrıntılı incelenmesine olanak sağlamıştır. Üçüncü aşama olan interpolasyon aşamasında öncelikle her sensör veri seti ayrı ayrı ele alınarak ayağın zeminle temas etmediği segmentler veri setlerinden çıkarılmıştır. Daha sonra, sensör verilerinin çözünürlüğü yirmi kat artırılacak şekilde interpolasyona tabi tutulmuştur. Bu titiz süreç, sadece verilerin hassasiyetini optimize etmekle kalmayıp, aynı zamanda yeterli örnek sayısına sahip kutucukları içeren histogram temsiline ulaşmayı da kolaylaştırmıştır. Veri setine yeni değerler eklenirken, verinin doğru temsilini koruma amacı güdülerek lineer interpolasyon yöntemi tercih edilmemiştir. Bunun yerine, verinin bütünlüğünü koruyarak daha doğru ve pürüzsüz bir temsil sağlaması nedeniyle kübik Hermite interpolasyon yöntemi benimsenmiştir. Bu yöntem, bireylerin yürüyüş desenlerinin çözünürlüğünü artırırken, daha gerçekçi bir veri setinin elde edilmesini mümkün kılmıştır. Dördüncü aşama olan verinin trendden arındırılması (Detrending) aşamasında, veri setlerinde yer alan trendlerin belirlenerek veri setlerinden çıkarılmasını amaçlanmıştır. Bu yöntem, trendden bağımsız olarak ortaya çıkan değişiklikleri daha hassas bir şekilde değerlendirmeye olanak tanır. Böylece, elde edilen sonuçların doğruluğunu artırarak daha güvenilir analiz sonuçlarına ulaşılmasına katkı sağlar. Bu araştırmanın önemli bir yeniliği, trend eğrisini oluşturan özel olarak tasarlanmış bir algoritmanın geliştirilmesidir. Verinin trendden arındırılması aşamasında, önerilen algoritma adım verilerinin eğilimini tespit eder. Bu algoritma, bir adıma ilişkin her veri noktasını bir alfa katsayısıyla ağırlıklı ortalama alarak bir önceki adıma ilişkin eğilim eğrisi ile birleştirir. Böylece bir trend eğrisi oluşturulur. Bu yöntem, ağırlıklı ortalama katsayısını iteratif olarak ayarlayarak hata miktarını minimize eden bir yaklaşım sunmaktadır. Belirli alfa ve hata eşik değerlerinin kesin bir biçimde belirlendiği bu yöntem, yürüyüş trendinin elde edilmesini ve bu trend etrafındaki veri dalgalanmasının ortaya çıkarılmasını sağlar. Bu yaklaşım, bireylerin yürüyüşlerindeki özgün dalgalanmaları yakalamayı ve yürüyüş verilerini trendlerinden arındırmayı hedefler. Beşinci aşamada, Tsallis entropisi değerleri hesaplamaları gerçekleştirilmiştir. Tsallis entropisi, entropi değerini belirleyen önemli bir parametre olan q'ya bağlıdır. Bu parametre, olasılık dağılımının geniş bir spektrumda ölçülmesine ve sistem özelinde daha hassas analizlere imkân tanır. Q parametresi, sistem özelliklerine göre ampirik yöntemlerle belirlenir ve bu çalışmada optimal q parametresi ampirik yöntemlerle 0.82 olarak belirlenmiştir. Tsallis entropisi hesaplamaları hem tüm yürüyüş için hem de her bir adım için ayrı ayrı gerçekleştirilmiştir. Bu yaklaşım, vestibüler sistem bozukluğu olan bireylerin yürüyüş dinamiklerinin daha kapsamlı değerlendirilmesini mümkün kılmıştır. Altıncı ve son aşamada kısa yürüyüş verilerinden elde edilen Tsallis entropisi değerleri kullanılarak sağlıklı ve hastalıklı bireyler arasında ayırt edici öznitelikler tespit edilmeye çalışılmıştır. Bu aşamanın soucunda makina öğrenimi için iki öznitelik seti elde edilmektedir. Bu setlerden biri bireylerin yürüşlerinin tamamından elde edilen her bir algılayıcı için Tsallis entropisi değerleridir; diğer öznitelik seti ise her bir adıma ait Tsallis entropisi değerlerinin sıfırdan sapma değerleridir.

Makine öğrenimi algoritmaları ile, kuvvet verilerinden elde edilen öznitelikler kullanılarak, bireyler “sağlıklı” ve “vestibüler sistem bozukluğu olan” kategorilerine sınıflandırılmıştır. SVM (Gaussian çekirdekli), KNN (Cosine) ve Lojistik regresyon, sırasıyla %95, %95 ve %93,3'lük başarı oranları ile en iyi performansı gösteren sınıflandırıcılar olarak tespit edilmiştir. Bu bulgular, Tsallis entropisi hesaplamaları ve öznitelik çıkarma yöntemlerinin, vestibüler sistem bozukluğu olan bireylerin tanısında etkili olduğunu göstermektedir. Sunulan metodoloji, vestibüler sistem bozukluklarının tanı ve izleme süreçlerinde klinik uygulamalar için potansiyel bir strateji teşkil etmektedir. Bu strateji, klinik değerlendirmelerin objektivitesini ve doğruluğunu optimize edebilir. Bu çalışma, Tsallis entropisi hesaplamalarının ve öznitelik çıkarma metodolojisinin, vestibüler sistem bozukluğu olan bireylerin tanısında değerli araçlar olabileceğini belirtmektedir. Bu yöntemlerin daha geniş ölçekli çalışmalarda incelenmesi, vestibüler rehabilitasyon pratiğine ve denge bozuklukları olan bireylerin yaşam kalitesine olan olası katkılarına daha fazla ışık tutabilir.

Çalışmanın bulguları, Tsallis entropisi hesaplamalarının ve öznitelik çıkarma yöntemlerinin yürüyüş verisi analizinde yeni bir yaklaşım sağladığını göstermektedir. Bu objektif değerlendirme aracı, vestibüler sistem bozukluğunun tanısında ve izlenmesinde sağlık profesyonellerine yardımcı olabilir, daha iyi tedavi ve yönetim stratejilerinin geliştirilmesine katkıda bulunabilir. Ayrıca, çalışma, yürüyüş verilerinin analizi üzerine yapılan araştırmalarında entropi tabanlı özniteliklerin vestibüler sistem bozukluklarının tanı ve değerlendirmesindeki önemini vurgulamaktadır. Bulgular, entropi tabanlı analizin çeşitli klinik ortamlarda daha yaygın ve detaylı bir şekilde keşfedilmesi ve uygulanması için fırsatlar sunmaktadır.

Bu çalışma ile elde edilen yöntemlerin klinik uygulamaya entegre edilmesi, sağlık profesyonellerinin yürüyüş örüntülerine ilişkin kantitatif ve objektif ölçümler elde etmelerine olanak tanıyabilir. Bu yaklaşım, vestibüler sistem bozukluklarının doğru tanısını kolaylaştırabilir, erken müdahale fırsatı sunabilir ve kişiselleştirilmiş tedavi planlarının oluşturulmasını teşvik edebilir. Ayrıca, entropi tabanlı özniteliklerin doğasının tanı sürecindeki öznel değerlendirmeleri tamamlayarak daha kapsamlı bir yürüyüş örüntüsü anlayışı sağlamasının yanı sıra tanıların genel güvenilirliğini artırması olasıdır. Bu çalışmada gözlenen yüksek performans, bu konuda daha fazla araştırma için bir itici güç olarak hizmet etmektedir. Bu araştırma, vestibüler sistemle ilişkili denge bozukluklarını anlama ve ele alma konusunda Tsallis entropisi analizinin potansiyelini göstererek, daha geniş bir projeye önemli bir katkı sunmaktadır.

1. INTRODUCTION

This thesis centers on the vestibular system, which plays a critical role in preserving balance and spatial orientation abilities in daily human life. The initial section provides an introduction to the fundamental functions of the vestibular system, its role in human life, and how disorders within this system are identified. Following this, an examination of the framework and contextual background will be undertaken. This section will outline the primary motivations of the thesis, the research problems, and how the objectives will be achieved. The subsequent literature review will delve into the current approaches to the diagnosis of vestibular system disorders, the methods employed, and the main challenges faced in this domain. Lastly, an evaluation will be made on the principal hypotheses of this study and how these hypotheses will be tested.

1.1 Vestibular System

The vestibular system, a complex structure located within the inner ear, serves as the body's primary system for detecting changes in motion and maintaining equilibrium. It consists of two main components: the semicircular canals, which detect rotational movements, and the otolithic organs, which sense linear accelerations. Together, these structures relay vital information about body movements and position to the brain. This information is then processed in conjunction with visual and proprioceptive inputs, allowing humans to maintain balance, coordinate head and eye movements, and navigate through their environment. A visual representation of these components and their dynamic responses can be seen in Figure 1.1. Malfunctions or disorders within the vestibular system can disrupt these vital processes, resulting in a range of symptoms that profoundly affect an individual's ability to perform everyday tasks and engage in social activities.

The vestibular system, essential for maintaining balance and spatial orientation in humans, plays a pivotal role in daily activities and overall quality of life [1]. Disorders affecting this system have significant implications and require accurate diagnosis and effective monitoring for appropriate treatment. These disorders can manifest as vertigo, dizziness, and imbalance, leading to increased risk of falls and reduced mobility. Early detection and intervention are essential in managing vestibular system disorders and improving overall patient outcomes. Analyzing gait data and extracting meaningful features from it is an essential tool for assessing human locomotion. This process is crucial in diagnosing and monitoring various medical conditions, such as vestibular system disorders [2]. By understanding the intricacies of changes in gait data within the context of vestibular system disorders, effective diagnostic and therapeutic approaches can be developed.

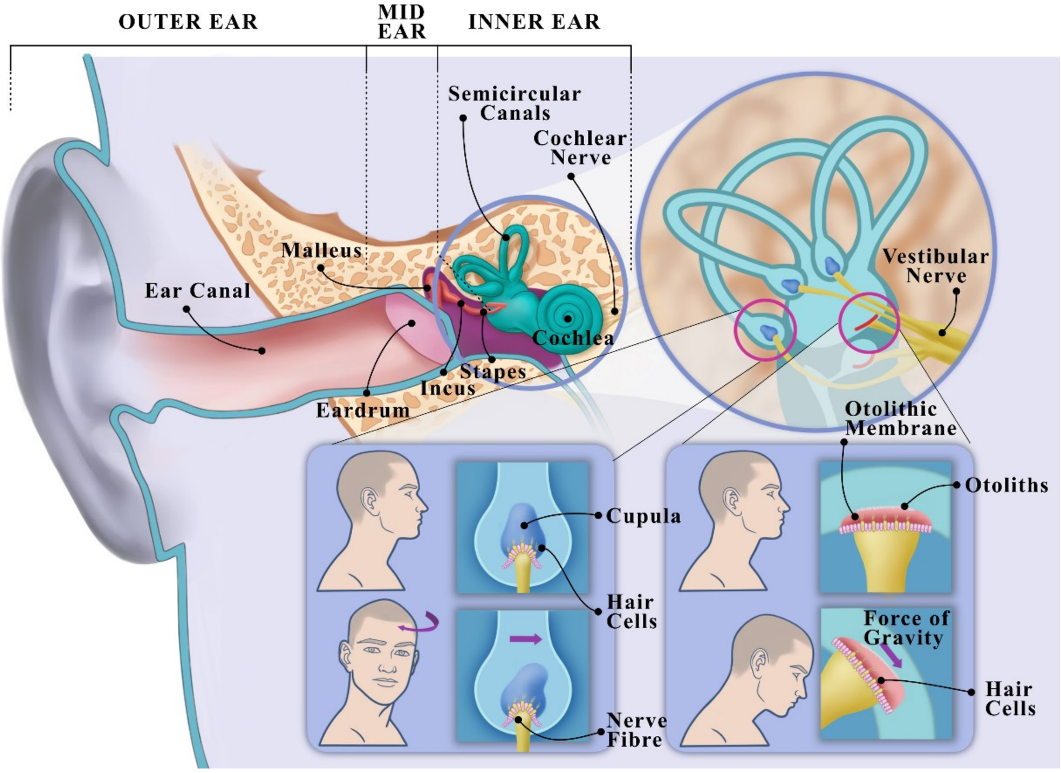


Figure 1.1: Visualization of the vestibular system components and their dynamic responses during individual motion.

Gait data analysis involves the systematic study of an individual's walking dynamics and patterns. This encompasses a detailed assessment of various attributes related to walking, such as force exertion, rhythmicity, speed, and consistency across distinct phases of the gait cycle. The gait cycle is intricately structured into two primary segments: the stance phase, during which the foot makes contact with the ground, and the swing phase, where the foot propels forward. Through comprehensive gait data analysis, one can identify anomalies and discrepancies that diverge from the standardized 'normal' gait parameters. Such variations may serve as markers for a range of health concerns, spanning musculoskeletal to neurological conditions.

1.2 Research Framework and Contextual Background

In the following sub-sections, the focus shifts to the foundation of this research. Initially, the background highlights the significance of the vestibular system in human balance and spatial orientation. Subsequently, the purpose, existing gaps in research, targeted objectives, research questions, and the significance of the study are systematically presented.

1.2.1 Contextual background

The background of this research lies in the significance of the vestibular system and its impact on balance and spatial orientation. Vestibular system disorders can result from various causes, including trauma, infections, or degenerative conditions [3]. These disorders can lead to symptoms such as dizziness, vertigo, and unsteadiness. Such symptoms can severely impair the quality of life, making even mundane tasks challenging for the affected individuals. The repercussions of the disease extend beyond medical implications to social and economic impacts. People with vestibular dysfunctions often face challenges in their daily activities, which leads to reduced work productivity and increased healthcare costs. Currently, the diagnosis of vestibular system disorders relies heavily on subjective assessments and clinical examinations. Therefore, there is a need to develop objective and reliable methods for diagnosis and monitoring.

1.2.2 Purpose of thesis

This thesis aims to employ contemporary classification methods and analyze force data from insoles to differentiate between healthy individuals and those with vestibular system disorders. A novel approach is introduced, leveraging Tsallis entropy as the key feature for analysis. The core objective is to explore the potential of gait data combined with Tsallis entropy-based methods for diagnosing vestibular system disorders. Through a detailed analysis of pressure patterns and extraction of relevant features from sensor data, this research seeks to unveil indicators that can assist in an objective assessment of vestibular system disorders. The outcomes of this study could provide pivotal insights for creating innovative diagnostic and therapeutic tools for individuals with VS disorders.

Traditional diagnostic methods for vestibular system disorders often involve time consuming procedures, subjective interpretations, or lengthy assessment processes. This study's approach hopes to overcome these challenges by using a non-invasive, objective, and efficient method, hence potentially revolutionizing the current diagnostic landscape.

1.2.3 Research gaps and rationale for the current study

Despite the importance of accurate diagnosis and monitoring of vestibular system disorders, there is a research gap in the development of objective assessment tools. Existing methods often rely on subjective assessments and qualitative observations, which can introduce variability and limitations in the diagnostic process [4, 5]. Furthermore, the intersection of advanced data processing with gait analysis for vestibular disorders is still nascent, presenting a clear research opportunity. This research aims to bridge this gap by exploring the potential of gait data analysis and data processing techniques in providing objective and quantitative measures for the diagnosis and monitoring of vestibular system disorders.

Building upon the findings from the literature review, several research gaps emerge, necessitating further investigation of vestibular system dysfunction-related balance disorders. Despite numerous studies in various areas of the medical field, the detection of vestibular disorders is an area that has not received sufficient attention yet. This study aims to address these gaps by utilizing Tsallis entropy calculations and feature extraction methods in the analysis of pressure data for individuals with vestibular system disorders. By leveraging the advantages of Tsallis entropy, this study seeks to develop a more comprehensive and objective approach for assessing gait patterns, enabling improved diagnosis, monitoring, and treatment evaluation in vestibular system disorders.

The primary objective is to achieve swift and accurate diagnosis of vestibular system disorders through the analysis of rapidly collected sensor data. The utilization of Tsallis entropy and advanced feature extraction methods is geared towards obtaining high accuracy in the diagnosis process. This overarching goal emphasizes the need for efficient and precise diagnostics using real-time sensor data, aligning with the ultimate aim of enhancing the speed and accuracy of vestibular system disorder diagnosis.

Addressing these gaps will not only advance the scientific understanding of vestibular system disorders but also offer tangible benefits to patients. Through enhanced diagnostic techniques, patients can expect more precise treatments, potentially leading to faster recovery times and improved overall quality of life. Moreover, the broader medical community can benefit from streamlined processes, reducing the burden on healthcare systems and professionals. The amalgamation of these research efforts contributes to the ongoing pursuit of effective and objective tools for the assessment and management of vestibular system disorders.

1.2.4 Research objectives

The primary objectives of this thesis are:

1. To rapidly acquire and accurately evaluate data for proper assessment of VS dysfunctions.
2. To investigate the use of force data and determine if the identified gait features from it can serve as reliable indicators for assessing, diagnosing, and monitoring individuals with vestibular system disorders.
3. To employ data processing techniques, including Tsallis entropy calculations, for feature extraction from force data obtained from insoles.
4. To identify potential classifiers for accurate classification of individuals with vestibular system disorders based on the extracted features.

By achieving these objectives, this research aims to contribute to the development of objective assessment tools and methods for diagnosing and monitoring vestibular system disorders.

These research objectives are further elucidated by the following supporting research questions:

1. What are the discernible differences in gait parameters and patterns between healthy and VS diseased individuals?
2. Is using exponential polynomials sufficient to obtain the trend of gait data, or is a more detailed technique needed?
3. Which classification algorithms show the best performance in distinguishing between healthy and VS diseased individuals based on the extracted features?

These research questions guide the investigation and analysis conducted in this thesis. By addressing these research questions, the study aims to shed light on the potential of force data analysis and advanced data processing techniques in the field of vestibular system disorders, ultimately improving diagnostic accuracy, treatment planning, and therapeutic interventions for affected individuals.

1.2.5 Significance of the study

This study holds considerable significance in its potential to introduce objective and quantitative tools for the diagnosis and monitoring of individuals with vestibular system disorders. By utilizing data processing techniques, the study aims to identify features that can differentiate between healthy individuals and those with vestibular system disorders. The development of reliable assessment tools can lead to improved accuracy in diagnosis, personalized treatment approaches, and enhanced monitoring of the progression of vestibular system disorders.

The significance of this study lies in its potential to lead to innovative therapeutic interventions for vestibular system disorders. By identifying and quantifying the specific features of these disorders, the study provides valuable insights into the underlying mechanisms and pathophysiology, benefiting researchers and clinicians.

Furthermore, by refining the diagnostic process, this approach could lead to substantial cost savings in the healthcare sector. Early and accurate diagnosis can reduce the need for repeated medical visits, decrease the chances of misdiagnosis, and eliminate the costs associated with unnecessary treatments or interventions. On a broader scale, this could translate to reduced economic burden on healthcare systems and improved patient outcomes.

1.3 Literature review

This literature review segment provides a comprehensive examination of the vestibular system disorder, shedding light on its clinical management and diagnostic processes. Initial sections delve into the vestibular system's intrinsic complexities, established methods of interpretation, and the analysis of gait data in relation to vestibular system disorders. Attention is also devoted to gait data analysis in the context of other diseases causing balance impairments. Subsequently, the focus shifts to Tsallis entropy, exploring its potential in gait signal analysis and its application in the biomedical signal analysis realm. Studies involving other entropy methods in gait data analysis are also highlighted. Lastly, the section concludes by identifying pertinent research gaps and underlining the rationale for the current study, which aspires to address these identified areas of concern.

In summary, this literature review sets the stage for subsequent research by presenting a comprehensive overview of the vestibular system and the significance of entropy methods, providing a strong foundation for the study's objectives and methodology.

1.3.1 Vestibular system disorder-related literature

In this section, emphasis is placed on the vestibular system and its associated disorders. The pathophysiology and clinical management of vestibular dysfunctions are initially highlighted. Methods and techniques currently utilized for diagnosis and monitoring are subsequently explored. A significant portion is devoted to the influence of vestibular system disorders on gait patterns and the application of gait data analysis in detecting other balance-impairing diseases. This segment aims to encapsulate the critical role of the vestibular system in maintaining balance and coordination.

1.3.1.1 Pathophysiology and clinical management

The examination of vestibular system disorders is a matter that requires significant emphasis. These disorders encompass a wide range of conditions that affect the inner ear and its connections to the brain, leading to impairments in balance and coordination [3].

Common vestibular system disorders include Meniere's disease, benign paroxysmal positional vertigo (BPPV), and vestibular neuritis. Understanding the pathophysiology and clinical manifestations of these disorders is crucial for comprehending their impact on gait patterns [6]. Furthermore, the literature review also explores the diagnostic methods and treatment options available for vestibular system disorders, highlighting the importance of early detection and appropriate management in improving patients' quality of life [3, 7].

Early diagnosis and intervention for these disorders are crucial, not only for mitigating the immediate symptoms but also for preventing potential long-term complications that can arise if left untreated, and delays in treatment can exacerbate the condition, making recovery more challenging and prolonged.

1.3.1.2 Existing methods and techniques

Various methods are employed in the literature to identify the specific VS problem while the most popular clinical method is still the computerized dynamic posturography (CPD) [8]. While CPD offers detailed insights, it may not be easily accessible in all clinical settings. On the other hand, wearable sensors provide the advantage of real-world data collection but may sometimes lack the precision of controlled, clinic-based tools. The state-of-the-art methods basically are based on utilizing classification techniques following a machine learning step where the features are extracted from gait data. The gait data are especially used to give information about balance disorder. Within this context, analysis of gait data has emerged as a valuable tool in the diagnosis and monitoring of balance disorder-causing diseases, providing objective measures to assess motor impairments associated with these conditions.

Traditional approaches include visual observation and subjective rating scales, which have inherent limitations in terms of objectivity and accuracy [9]. However, advancements in sensor technology and computational methods have led to the development of more sophisticated tools, such as wearable sensors and computerized algorithms, enabling quantitative and objective analysis of gait data.

The utilization of wearable sensors, particularly insole sensors for analyzing force data, facilitates the collection of gait data in individuals' daily lives, eliminating the need for clinical environments. This approach allows data to be obtained in daily life, helping the patient avoid the stress of the clinical environment and potentially improving the accuracy of the diagnosis [10, 11]. Conversely, Individuals suffering from balance disorders may exhibit abnormal gait patterns due to the awareness of being observed, leading to deviations from the normal performance [12]. The analysis of force values on insole sensors can provide insights into changes in foot behavior during walking, enhancing accuracy and reducing stress associated with the diagnostic process.

Despite the progress made in analysis of gait data techniques, there are still limitations that need to be addressed. Current approaches often focus on a limited set of gait parameters and may not fully capture the complex nature of gait disturbances in vestibular system disorders. Furthermore, the lack of standardized protocols and the influence of confounding factors pose challenges in interpreting and comparing gait data across different studies.

1.3.1.3 Analysis of gait data in balance-impairing diseases

Analysis of gait data has emerged as a valuable tool in the diagnosis and monitoring of balance disorder-causing diseases, providing objective measures to assess motor impairments associated with these conditions. It has been extensively utilized in the evaluation of diseases such as Parkinson's disease (PD), Huntington's disease (HD), amyotrophic lateral sclerosis (ALS), and other related disorders. Numerous studies have demonstrated the effectiveness of analyzing gait data in identifying disease-specific gait abnormalities and distinguishing between different balance disorder-causing diseases. For instance, Nir Giladi et al. proposes a new clinical classification scheme for gait and posture and discusses the use of analysis of gait data in identifying disease-specific gait abnormalities [13]. As another example, Bovonsunthonchai et al. investigate the use of spatiotemporal gait variables in distinguishing between three cognitive status groups and discusses the potential of gait data analysis as a tool for early detection of imbalance-causing neurodegenerative conditions [14]. And, Guo Yao et al. summarizes researche on the effectiveness and accuracy of different gait related data analysis systems and machine learning algorithms in detecting Parkinson's disease [15].

Analyzing data gathered during walking plays a crucial role in various fields, including medical diagnostics and healthcare. In a study conducted by Ikizoğlu and Heyderov [16], they explore the significance of features extracted from IMU-sensor based data to diagnose vestibular system disorders. By examining the data collected during walking, they aim to identify patterns or abnormalities that can help in the early detection and treatment of such disorders.

Research by Agrawal et al. [17] emphasizes the importance of analyzing data gathered during walking to predict fall risks. They utilize wireless pressure sensors embedded in insoles and employ machine learning models to analyze the collected data. Through this approach, they are able to assess the likelihood of falls and provide early warnings or interventions, thus improving the safety and well-being of individuals at risk.

Furthermore, Bustamante et al. [18] highlight the significance of monitoring and assessing imbalance-causing neurodegenerative diseases using a portable wireless pressure sensing device. By analyzing the data obtained during walking, they can track changes in pressure patterns, gait abnormalities, and other relevant metrics. This information aids in the early detection, evaluation, and management of imbalance-causing diseases like neurodegenerative and VS-related diseases, ultimately improving the quality of life for patients.

In conclusion, the analysis of data gathered during walking plays a pivotal role in various areas, such as medical diagnostics, fall risk prediction, and monitoring of imbalance-causing diseases. By leveraging advanced technologies and machine learning models, researchers and healthcare professionals can extract valuable insights from this data, leading to improved diagnosis, prevention, and treatment strategies.

1.3.1.4 Analysis of gait data in vestibular system disorders

Analysis of gait data provides clinicians with valuable information about the functional limitations and compensatory mechanisms employed by individuals with vestibular system disorders [19]. By examining gait patterns, healthcare professionals can tailor treatment strategies and evaluate the effectiveness of interventions.

The subsequent examination primarily concentrates on the analysis of human locomotion, specifically within the framework of vestibular system disorders. Analysis of gait data yields significant revelations regarding the functional limitations induced by these disorders, thereby facilitating their diagnosis, treatment, and monitoring. Gait parameters, including stride length, step width, and cadence, frequently exhibit alterations among individuals afflicted with vestibular system disorders [20]. The quantitative analysis of such parameters enables an objective evaluation of the disorder's severity and facilitates the tracking of temporal changes.

In the realm of vestibular system disorders, gait data provides a methodical exploration of an individual's locomotion characteristics. This study delves deep into the nuances of walking, examining elements like force application, walking rhythm, speed, and the uniformity observed during the gait cycle's distinct phases. This cycle is primarily split into two segments: the stance phase, where the foot is grounded, and the swing phase, marking the foot's forward movement. An in-depth analysis of gait data can pinpoint deviations from what's deemed as the 'normal' walking parameters. These deviations potentially act as indicators of various health issues, from musculoskeletal anomalies to neurological disorders.

By examining gait parameters, insights have been gained into the impact of vestibular-related impairments on balance control. This comprehensive analysis aids in early detection, accurate diagnosis, and monitoring of these disorders. For example, studies by Ratan Das et al. [21] provides a comprehensive review of the literature and discusses how analysis of gait data can be used as a clinical tool to better diagnose and manage gait and balance impairments. As another example, A. R. Wagner et al. [22] discusses how analysis of gait data can be used to assess vestibular-related impairments in older adults, and how these impairments can impact balance control. In a study by Schmidheiny et al., the team delved into the discriminant validity and test-retest reproducibility of gait assessments in patients with vestibular dysfunction [23]. This research underscores the significance of an objective evaluation of gait in individuals with vestibular disturbances.

1.3.2 Tsallis entropy-related literature

In this section, the spotlight is turned to Tsallis entropy and its applications within biomedical signal analysis. Initially, an overview is provided on the role of Tsallis entropy in gait signal analysis. This is followed by an exploration into its broader applications in the realm of biomedical signals. Lastly, the section delves into how gait data analysis has been approached using other entropy methods, offering a comparative perspective. The aim of this segment is to delineate the emerging prominence of Tsallis entropy in advancing diagnostic techniques, especially in the context of gait data.

1.3.2.1 An overview of Tsallis entropy's role in gait signal analysis

Tsallis entropy offers a flexible framework for signal analysis, allowing the adjustment of a parameter to control the emphasis on rare events or outliers. This adaptability makes Tsallis entropy well-suited for capturing the intricate and non-linear nature of gait signals affected by vestibular system disorders. In this context, Tsallis entropy is an adaptable tool for analyzing gait signals, offering insights into their non-linear nature and enabling the development of effective diagnostic and therapeutic approaches.

Tsallis entropy has proven to be effective in diverse domains such as physics, information theory, and economics, enabling a more comprehensive comprehension of systems with long-range correlations and heavy-tailed distributions [24]. In this context, Tsallis entropy can capture long-range correlations within the data, offering a richer and more nuanced understanding of gait dynamics.

1.3.2.2 Tsallis entropy in biomedical signal analysis

Based on prior research, this section examines the varied implementations of Tsallis entropy in the analysis of biomedical signals. It emphasizes the effective utilization of Tsallis entropy in quantifying intricacy, identifying anomalies, and discerning distinct physiological conditions in diverse biomedical signals. The efficacy of Tsallis entropy in offering fresh perspectives on analysis of gait data in vestibular system disorders is apparent based on its successful utilization in alternative fields.

As an example of Applications of Tsallis entropy in biomedical signal analysis, Al-Nuaimi et al. examine the application of Tsallis entropy as a promising approach for assessing alterations in EEG signals among individuals diagnosed with dementia. Their study investigates the capacity of Tsallis entropy to gauge variations in the intricacy of EEG signals, emphasizing its potential as a biomarker for dementia [25]. As another example, Gao et al., in their publication, present a thorough examination of Tsallis entropy-derived metrics in the context of biomedical signal analysis. They explore the capacity of Tsallis entropy to effectively quantify the generation rate of meaningful information within a dynamic system, while also highlighting its potential applications across diverse biomedical signal analysis endeavors [26].

The use of Tsallis entropy in biomedical signal analysis underscores the ongoing evolution of diagnostics, where computational tools are playing an increasingly pivotal role. Leveraging such techniques can potentially revolutionize the accuracy and efficiency of medical diagnoses, leading to more timely and targeted interventions.

1.3.2.3 Analysis of gait data with other entropy methods

This section reviews previous studies that have investigated the use of entropy methods in assessment of locomotion data, specifically in the context of imbalance-causing diseases such as neurodegenerative and VS-related diseases. These studies have demonstrated the effectiveness of entropy methods in capturing subtle changes in gait patterns and differentiating between healthy and diseased individuals. The findings from these studies provide a strong rationale for further exploring the application of Tsallis entropy in the current study.

In their study, Kim et al. explore the utilization of analysis of gait data and machine learning techniques in diagnosing vestibular neuritis. The findings of these studies indicate that the application of multiscale approximate entropy (MAE) method and other gait analysis techniques holds promise for facilitating the diagnosis and treatment of neurodegenerative diseases [27]. As another example, Haid et al., in their article, investigate the utilization of sample entropy (SaEn) for the analysis of postural control among individuals diagnosed with Parkinson's disease. Their study aims to test the hypothesis that sample entropy correlates with the complexity of postural movement patterns. The findings demonstrate that sample entropy can successfully differentiate between healthy individuals and those affected by Parkinson's disease [28].

1.4 Hypothesis

It is postulated that there are discernible differences in gait parameters and patterns between healthy individuals and those diagnosed with vestibular system disorders, which can be captured and quantified through the application of advanced data processing techniques [12, 29]. In this context, it is hypothesized that Tsallis entropy calculations, as a novel data processing method, can provide valuable insights into the complexity and variability of gait patterns in individuals with vestibular system disorders, further enhancing the differentiation between healthy individuals and those with the disorder.

Based on preliminary evidence suggesting a link between vestibular system disorders and alterations in gait patterns, it is postulated that Tsallis entropy calculations can effectively quantify these differences. Specifically, we hypothesize that this method will reveal discernible variations in gait parameters between healthy individuals and those with vestibular system disorders.

1.4.1 Formulation of hypothesis

The hypothesis suggests distinguishable differences in force data and patterns between healthy and VS diseased individuals. These differences can be captured and quantified through the application of advanced data processing techniques, specifically Tsallis entropy calculations.

As already mentioned in Section 1.1, Vestibular system disorders have a significant impact on an individual's balance, coordination, and overall gait performance. Existing literature indicates that the analysis of biomechanical data, such as force patterns during walking, can play a crucial role in identifying alterations in gait. This study proposes that when combined with Tsallis entropy calculations, this data might offer a more comprehensive insight into the gait patterns associated with vestibular system disorders. Building upon the existing body of literature, the hypothesis for this study is articulated as follows:

"The utilization of Tsallis entropy calculations, as a novel data processing method, can provide valuable insights into the complexity and variability of gait patterns in individuals with vestibular system disorders, enhancing the differentiation between healthy individuals and those with disorder."

1.4.2 Supporting arguments for the hypothesis

As presented in the Results section, various aspects of this study highlight a connection between gait data analysis and vestibular system disorders. Observations suggest that the histogram of preprocessed sensor data contains more bins for individuals with vestibular system disorders. This increase in the number of bins is attributed to compromised balance and coordination, leading to greater fluctuation around the trend curve.

Furthermore, Tsallis entropy calculations and feature extraction methods appear promising in emphasizing the subtle alterations in gait patterns associated with vestibular system disorders. By analyzing these entropy levels and the detailed step-wise entropy changes, the complexity and variability of gait patterns can be quantified, offering a comprehensive and objective assessment of the disorder.

The objective of this research is to deepen the understanding of vestibular system disorders and to introduce a potentially significant diagnostic tool. By utilizing gait parameter analysis combined with advanced data processing techniques, this study aims to pave the way for improved diagnosis, monitoring, and treatment strategies for individuals affected by vestibular system disorders.

2. TSALLIS ENTROPY METHOD

Tsallis entropy, named after the physicist Constantino Tsallis, is a generalized form of entropy that has found applications in various fields, including biomedical research. In the context of biomedicine, the Tsallis entropy method has proven to be a valuable tool for analyzing complex systems and understanding the dynamics of biological processes.

2.1 Entropy

Entropy is a property primarily used to measure the disorder or randomness within a dynamic system. The well-known Shannon entropy (SE), formulated as equation 2.1, is based on Boltzmann-Gibbs statistical mechanics and is capable of describing the structure of extensive systems with short-term microscopic correlations [30, 31].

$$SE = -k_B \sum_{i=1}^{N} p_i ln(p_i) \tag{2.1}$$

In equation 2.1, k_B represents the Boltzmann constant, a fundamental physical constant that delineates the relationship between temperature and energy. Historically, the Boltzmann constant emerged as a proportionality factor in the microscopic description of ideal gas laws, and it plays a pivotal role in bridging microscopic and macroscopic views in thermodynamics. Specifically, it scales the average kinetic energy of particles to the thermodynamic temperature of a system.

However, in numerous analytical situations, especially within information theory and certain statistical mechanics contexts, it's conventional to normalize the Boltzmann constant to unity. This normalization simplifies calculations without sacrificing the conceptual underpinnings of the entropy measure.

$$SE' = -\sum_{i=1}^{N} p_i ln(p_i) \tag{2.2}$$

Equation 2.2 embodies the Shannon entropy in scenarios where the Boltzmann constant k_B is taken as unity. In this framework, N is the total number of microstates and p_i delineates the probability associated with the i-th microstate.

For systems with long-term interactions, or systems that exhibit long-term memory effects, the effectiveness of applying SE for the aforementioned purpose decreases [32]. This is particularly true when dealing with complex systems that involve intricate relationships and dependencies over extended periods. In such cases, a more comprehensive approach is needed to capture the underlying dynamics and extract meaningful information from the time series data. This is where the generalized structure of Boltzmann-Gibbs statistics comes into play, incorporating the concept of Tsallis entropy (TE) within the framework of non-extensive statistics [33]. By considering the TE, we can delve deeper into the hidden information embedded in the time series, unlocking valuable insights that might be overlooked by traditional methods. The TE provides a powerful tool to explore the complexities and intricacies of these long-term interacting systems, enabling a more nuanced understanding of their behavior and uncovering previously undiscovered patterns. Therefore, incorporating TE within the analysis of time series data proves to be a crucial step towards unraveling the full potential and richness of information contained within these intricate systems.

2.2 Tsallis Entropy

The Tsallis entropy formula provides a powerful tool for quantifying the complexity and structure of given dataset [11]. The parameter q in the Tsallis entropy formula represents a dimensionless entropic index, adjusting the entropy metric to capture specific features inherent in the analyzed dataset. By adjusting the value of q, the entropy metric can be tailored to capture particular features inherent in the analyzed dataset. The Tsallis entropy, denoted as S_q, is a generalization of the Shannon entropy and is defined as equation 2.3 [34].

$$TE = S_q = \frac{k_B}{q-1}\left(1 - \sum_{i=1}^{W} p_i^q\right) \tag{2.3}$$

Here, S_q represents the Tsallis entropy, k_B is the Boltzmann constant, q serves as a generalization parameter. Additionally, W signifies the total number of possible states, while p_i corresponds to the probability of the i-th state.

In the context of Tsallis entropy, as in Shannon entropy, the Boltzmann constant is frequently introduced. However, in disciplines such as information theory or data analysis, the primary interest often lies in the relative behavior of the system rather than its absolute scale. Therefore, it's customary to normalize k_B to unity to simplify analytical and computational processes.

Given this normalization, the Tsallis entropy is expressed as:

$$TE = S_q = \frac{1}{q-1}\left(1 - \sum_{i=1}^{W} p_i^q\right) \tag{2.4}$$

By setting $k_B = 1$, the role of the entropy in measuring the intrinsic properties of the data is emphasized, abstracting from the physical dimensions that k_B might introduce. This dimensionless representation of entropy is found to be especially pertinent when comparative analyses across diverse datasets are conducted.

In equation 2.4, q $(q\epsilon\mathcal{R})$ can be considered as a parameter to indicate the degree of non-additivity. This is because, in the context of two independent systems X and Y as represented by equation 2.5, $(1-q)$ serves as a quantification of the deviation from additivity.

$$TE(X+Y) = TE(X) + TE(Y) + (1-q) * TE(X) * TE(Y) \tag{2.5}$$

In the calculation of Tsallis entropy, the selection of parameter q is based on the specific characteristics of the dataset being analyzed. The value of q determines the non-extensive behavior of the entropy and can be chosen accordingly, taking into account the properties of the dataset [35].

Unlike Shannon entropy, which corresponds to $q = 1$, Tsallis entropy allows for a more flexible representation of information content. By adjusting the value of q, researchers can explore and quantify different aspects of the data distribution. For $q < 1$, Tsallis entropy emphasizes the contribution of rare events, making it more suitable for capturing long-tail or heavy-tailed distributions. Conversely, for $q > 1$, Tsallis entropy focuses on the more frequent events, which is useful for capturing data distributions with prominent peaks or clusters [36, 37]. Briefly, while $q > 1$ pertains to sub-extensive statistics, $q < 1$ aligns with super-extensive statistics.

Furthermore, the Tsallis entropy formula is widely used in physics, information theory, and data analysis. Its flexibility and adaptability make it a valuable tool for understanding complex systems. Tsallis entropy considers the non-extensive behavior of entropy, providing a comprehensive perspective on data information content and distribution. Its versatility enables the extraction of meaningful insights and the discovery of hidden patterns and structures from diverse datasets.

2.3 Application of Tsallis Entropy in Data Analysis

The Tsallis entropy method has found applications in various fields, including data analysis and pattern recognition. In the analysis of gait parameters, Tsallis entropy provides a measure of the complexity and irregularity of gait patterns. By quantifying the information content in gait data, it allows for the extraction of meaningful features related to the dynamics of human locomotion.

In gait data analysis, Tsallis entropy is often calculated based on the histograms of specific gait parameters, such as force amplitude, step length, step duration, or joint angles. By examining the distribution of these parameters and applying the Tsallis entropy formula, researchers can obtain a measure of the system's complexity and irregularity. This measure can then be used to compare gait patterns between different individuals, groups, or conditions.

As an example, application of Tsallis entropy in data analysis, in addition to examples in Section 1.3.2.2, Zhang et al. propose a measure called Tsallis entropy area (TsEnA) to quantify burst suppression (BS) activity in EEG data after brain injury, demonstrating its correlation with neurological deficit scores and suggesting its potential as a clinical tool for estimating the severity of brain damage following cardiac arrest [38]. Again, Tong et al use TE of EEG signals as a measure of brain injury in their study [37]. As another example, Thilagaraj et al. aimed to classify epileptic seizures using EEG data segments from the University of Bonn database. They introduced a novel feature related to Tsallis entropy and utilized five different classifiers. Their method achieved high accuracy ranging from 92.67% to 100% using a Decision tree classifier, while also having the fastest computation time compared to other features in the literature. The authors proposed that their method can be easily installed as a software tool and holds potential for real-time detection and prediction of epileptic seizures [39]. Moreover, Redelico et al. compare different permutation entropies as classifiers for EEG records of normal and pre-ictal states. Symbolization techniques are used to derive discrete probability distribution functions, and the entropies are used as independent variables in logistic regression models. All permutation entropies perform well, with high accuracy and sensitivity, suggesting their potential for automatic EEG signal classification [40].

2.4 Advantages of Tsallis Entropy in Analysis of Gait Parameters

The use of Tsallis entropy in gait data analysis offers several advantages. First, it can capture non-linear and non-Gaussian characteristics of gait data, which are often present in complex systems such as human movement. This enables a more accurate representation of gait dynamics compared to traditional entropy measures.

Additionally, Tsallis entropy allows for the customization of the entropy calculation through the parameter q. By adjusting q, the method can adapt to different types of gait patterns and provide insights into specific aspects of gait dynamics, such as regularity, variability, or asymmetry. This flexibility makes it a valuable tool for analyzing and comparing gait patterns in individuals with different conditions or disorders.

Furthermore, Tsallis entropy provides a robust and efficient method for feature extraction in analysis of gait data. By quantifying the complexity and irregularity of gait patterns, Tsallis entropy can identify distinctive features that differentiate individuals or groups. These features can then be used in machine learning algorithms or classification models to develop diagnostic or monitoring tools for gait-related disorders.

One of the main advantages of the Tsallis entropy method is its ability to capture the non-linear and long-range dependencies present in biological systems [32, 41]. Traditional entropy measures, such as Shannon entropy, assume that events are independent and identically distributed, which is not always the case in biological systems where interactions and correlations play crucial roles. Tsallis entropy, with its parameter q, allows for a more flexible characterization of these complex interactions. In biomedical research, the Tsallis entropy method has been applied to a wide range of applications. One such application is in the analysis of electroencephalogram (EEG) signals [39]. EEG is a non-invasive technique used to measure brain activity, and the analysis of EEG signals can provide valuable insights into neurological disorders and cognitive processes. By applying the Tsallis entropy method to EEG data, researchers have been able to identify abnormal brain activity patterns and distinguish between different brain states with higher accuracy than traditional entropy measures [37, 40].

Overall, the Tsallis entropy method has proven to be a valuable tool in biomedical research, enabling a more comprehensive analysis of complex biological systems. Its ability to capture non-linear and long-range dependencies provides researchers with a more accurate representation of biological processes and opens up new avenues for understanding and addressing various biomedical challenges.

In conclusion, the Tsallis entropy method is a statistical approach that extends Shannon entropy, providing a customized measure of information content in complex systems like analysis of gait data. Its application in gait data analysis offers advantages in capturing non-linear dynamics and providing customizable insights into gait patterns. By utilizing Tsallis entropy, researchers can gain valuable insights into the characteristics of gait and its variations in individuals with different conditions.

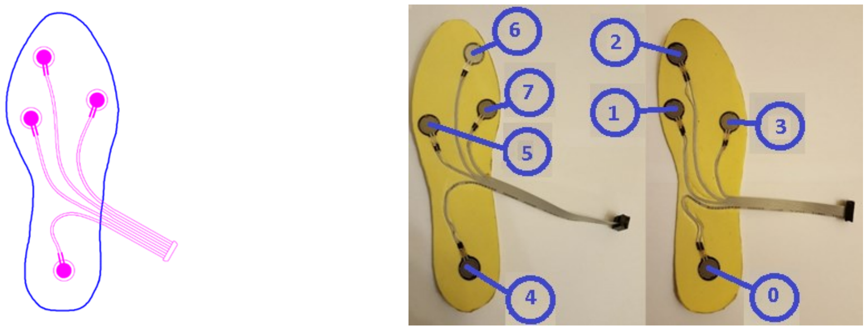
3. DATA ACQUISITION AND PROCESSING

3.1 Data Acquisition

The examination of gait data analysis studies in the literature reveals that the distribution of weight on the soles of the feet has a significant impact at four main points, as depicted in Figure 3.1a [42-46]. The selection of sensor placement mentioned above was determined based on previous studies.

3.1.1 Sensor placement

To ensure data collection without disturbing the natural walking patterns of the participants, five pairs of insoles with different sizes (36, 38, 40, 42, 44 - according to European Standards) were manufactured, and the sensors were placed in appropriate positions. Prior to the commencement of the experiment, the correctly sized insoles were inserted into the subjects' shoes. For the production of the insoles, a durable and soft plastic material commonly employed in the manufacturing of orthopedic products was utilized.



(a) (b)

Figure 3.1: (a) Sensor placement on the insole. **(b)** The numbering of the sensors S0 to S7 (Top view) [47].

Force-sensitive resistors (FSR) were chosen as force sensors, as they are widely used in gait data analysis applications and offer several advantages [48]. Specifically, the FSR402-Short tail model from Interlink was selected as it is highly suitable for placement within insoles due to its physical dimensions. Additionally, this sensor demonstrates acceptable repeatability error [49]. The characteristics of the sensor can be found in Table 3.1, and the numbering of the sensors on the insoles, from S0 to S7, is illustrated in figure 3.1b.

Table 3.1: Characteristics of The Sensor Fsr402-Short Tail

PARAMETER	VALUE
Operation Range	0.2N-20N
Physical	fpad18.3mm, fsens12.7mm
Dimensions	thickness 0.46mm
Repeatability	±0,02
Idle Resistance	>10MΩ
Hysteresis	10 % max.
Rising Time	<3 µseconds

The sensor referenced in our study, as delineated in Table 3.1, is the FSR402-Short Tail. It is capable of detecting forces ranging from as light as 0.2 Newtons to as robust as 20 Newtons, within an operational range of 0.2N-20N. Regarding its physical dimensions, the outermost layer or pad of the sensor boasts a diameter of 18.3mm, while its force-sensitive area measures 12.7mm in diameter, and the entire unit maintains a slim profile at a mere 0.46mm thickness. Repeatability, indicating the sensor's consistency in readings, suggests that measurements under identical conditions can differ by a maximum of 2% from their mean, underscoring its reliability. Its idle resistance, denoting the resistance when devoid of force, exceeds 10 Megaohms. Another critical attribute, hysteresis, signals a potential 10% variation in output based on the force's increasing or decreasing direction. The rise time, which demarcates the time the sensor requires to alter its output from 10% to 90% of the end value, underscores its brisk responsiveness with a rate of less than 3 microseconds. Collectively, these specifications provide insight into the sensor's functionality and capacity, ensuring a comprehensive grasp of its pivotal role in our research.

3.1.2 Data acquisition procedure

The data collection was carried out in the clinical setting of the Audiology Department at Cerrahpaşa Medical School, Istanbul University in Istanbul, Türkiye. The process was conducted in compliance with the principles outlined in the Helsinki Declaration. Prior to the start of the process, approval was obtained from the Istanbul University

Ethics Committee (Approval number: A-57/07.07.2015). In addition, informed consent was taken from all subjects to participate in the study. For individuals with VS problems, their conditions had already been diagnosed by the audiologists using conventional systems, such as Computerized Dynamic Posturography.

Data collection occurred on weekends to minimize subject stress and prevent interference from other devices in the environment. The subjects walked along a 12-meter track inside the clinic, and they were asked to walk the path twice. The first walk aimed to help them become familiar with the environment and reduce any possible stress, while the data from the second walk were used for analysis in general. In some cases, subjects walked a third time when needed as a result of the audiologists' observations. The data collection was optimized to a short span of 10-15 seconds, which is considerably shorter than most experiments in the literature.

3.1.3 Data acquisition system

The gait data was collected using insole force sensors, offering a detailed and reliable source of information. The force sensor data were initially captured by an Arduino Mega unit carried by each subject and then wirelessly transferred to a nearby laptop via an HC-06 Bluetooth module. Sampling was conducted at a rate of twenty samples per second from each sensor, simultaneously from all sensors.

In order to convert the force to voltage, a voltage divider was created using a 1k ohm resistor connected in series with the force sensor. The voltage divider was supplied with a 5V DC input. While the force sensor utilized in the study demonstrated high repeatability, its force-resistance characteristic was non-linear. Consequently, the device underwent calibration in the laboratory using known weights. As a result, the equation of the regression curve was obtained as equation 3.1.

$$w = e^{\frac{v_o + 0.2245}{0.9265}} \tag{3.1}$$

Here, w represents the relationship between the weight/force in Newtons and v_o is the output voltage of the voltage divider in volts. Ten percent deviation from the values obtained by equation 3.1 was taken as the criterion that would require the relevant sensor not to be used in the experiments.

The calibration process was performed individually for each sensor utilized in the experiments. This individual calibration ensured that each sensor operated at its optimum performance level, accounting for any inherent variances or manufacturing discrepancies. As a result, we were able to maintain consistency in data acquisition, thereby minimizing potential errors and achieving more reliable results.

3.1.4 Participants' information

Regarding the subjects who participated in the study, detailed information is provided in Table 3.2.

Table 3.2: Information About the Subjects

Physical	Healthy (30)		Diseased (30)	
Characteristics	Male (15)	Female (15)	Male (13)	Female (17)
Age	54,3±8,5	55,1±7,9	54,5±8,5	55,7±8,4
Mass [kg]	66,6±9,8	65,1±8,8	65,9±10,2	63,9±8,6
Height [cm]	169,2±10,0	164,0±6,2	170,3±8,8	162,7±6,3

The distribution of diseases among the subjects' experiencing discomfort was analyzed using computerized dynamic posturography, which was conducted under the supervision of audiologists. The distribution of the subjects whose specific disease was detected by computerized dynamic posturography by audiologists is given in Table 3.3. The data obtained through this rigorous methodology provides valuable insights into the distribution and incidence rates of these diseases, contributing to our understanding of their impact on postural stability.

Table 3.3: The Distribution of The Diseases of Suffering Subjects

Vestibular Disorders	Male	Female
BPPV*	6	8
UVW*	3	4
Meniere	3	3
Vestibular Neuritis	1	3

* BPPV- Benign paroxysmal positional vertigo, UVW-Unilateral vestibular weakness.

3.1.5 Ethical considerations

In order to uphold the confidentiality and privacy of all participants, their identities have been anonymized for the purpose of this article. This precautionary measure is implemented to protect the individuals involved and to maintain the integrity of the study's findings. By anonymizing the participants' identities, their personal information and any potential identifying details have been removed, ensuring their anonymity throughout the publication process.

This ethical practice serves to safeguard the rights and well-being of the participants, aligning with the principles of research ethics. By maintaining confidentiality, the study promotes a secure environment for participants to share their experiences and data without fear of retribution or breach of privacy. Anonymization also contributes to the overall reliability and credibility of the research, as it allows for unbiased analysis and interpretation of the collected information.

3.2 Data Processing

In the field of data analysis and signal processing, the preprocessing of captured data plays a crucial role in extracting meaningful insights and information. This is particularly important in complex domains such as analysis of gait data, where raw sensor data poses challenges and limitations when directly applying analytical methods [50]. Therefore, in order to enhance the comprehensibility and effectiveness of the data, a well-defined data processing pipeline is employed. This section introduces the key stages involved in the data processing approach used in the thesis, aiming to make the data more concise and suitable for further analysis.

Therefore, in order to make the data more concise and comprehensible, the application process of the method has been divided into six stages.

3.2.1 Stage 1 – normalization

The first stage in the data processing pipeline is normalization, where the captured raw data is transformed to a common scale for each subject. This step ensures that the data is comparable across different individuals and allows for meaningful comparisons to be made. Additionally, during normalization, the regions of interest associated with the portions where walking occurs are identified and framed., while irrelevant areas are removed from the dataset. This initial processing step is crucial for focusing the analysis of gait data on the parts where gait occurs and excluding incomplete or inaccurate data from the overall analysis.

Data are normalized for each subject to a range of 0 – 1. Then, the parts suitable for the data analysis, in other words, the parts where the gait takes place are framed and marked in the datasets.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{3.2}$$

Here, X is the original/raw data, X_{norm} stands for the normalized data. X_{min} and X_{max} represent minimum and maximum values, respectively.

Subsequently, the data was visualized. The areas suitable for data analysis in the images, i.e., the regions where the gait took place, were outlined and the remaining areas were removed from the dataset. These processes are visualized for a sample subject in figure 3.2. During this stage, data corresponding to the first and last steps were excluded from the overall gait data. These steps were omitted as they do not provide accurate information due to their incomplete dynamic behavior.

The normalization stage is followed by the step analysis phase, where the data undergoes further refinement. In this phase, the captured data is examined in detail to identify the regions that correspond to individual steps within the gait cycle. By accurately delineating each step, the subsequent analysis can focus on the relevant gait parameters and discard any incomplete or erroneous data associated with the transitional phases between steps. Removing the data from the first and last steps, which exhibit incomplete dynamic behavior, ensures that the subsequent analysis is based on reliable and consistent gait patterns. The step analysis stage enhances the overall quality of the gait data and contributes to a more precise and accurate assessment of individuals with vestibular system disorder.

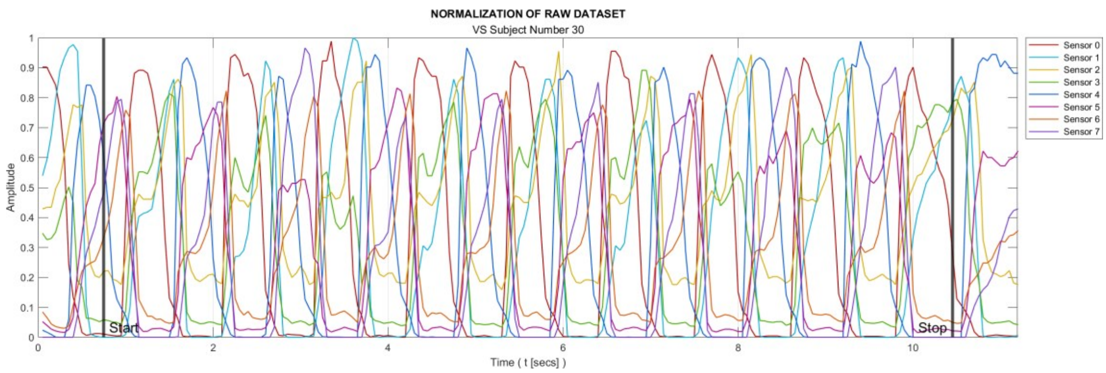


Figure 3.2: Normalization of raw sensor data and the framing of suitable segments for data analysis.

3.2.2 Stage 2 – step analysis

The subsequent stage is step analysis, which involves analyzing the data from each sensor separately for the subject's right and left feet. Algorithms are employed to detect the contact and lift-off phases of the foot during walking. The data from each sensor is segmented into parts corresponding to walking steps, enabling a more detailed analysis of the gait pattern.

For each subject, the data from each sensor were analyzed separately according to the data obtained from the subject's right and left feet. The areas where the foot makes contact with the ground and where the foot is lifted off the ground are detected by an algorithm.

The next stage involves analyzing data from each sensor separately for the subject's right and left feet. Algorithms detect the contact and lift-off phases of the foot during walking, enabling a detailed analysis of the gait pattern. Each sensor's data is segmented into walking steps for a more comprehensive examination. The dataset of a subject, consisting of eight lines, is decomposed into four-line datasets for the right and left feet. The data is then combined with the maximum operation by calculating the maximum values for each column of this four-row dataset. The combination of the data obtained from a subject's feet using the maximum operator can be represented by equation 3.3.

$$\begin{aligned} S_{Rmax} &= \max(S_0, S_1, S_2, S_3) \\ S_{Lmax} &= \max(S_4, S_5, S_6, S_7) \end{aligned} \tag{3.3}$$

Here, S_{Rmax}, S_{Lmax} stand for the time-series data that contains the maximum values of each column of the sensors data of the right (S_{Rmax}) and left (S_{Lmax}) foot. S_0, S_1 ... S_6, S_7 represent the time series data of the sensors on the insoles.

By calculating the differential of the datasets obtained from equation 3.3, it is possible to detect and cluster the steps of the subject's right foot. Subsequently, the same operations are applied to the sensor data obtained from the left foot. These processes enable the data from each sensor to be segmented into parts according to walking steps. These processes are visualized for a sample subject in figure 3.3.

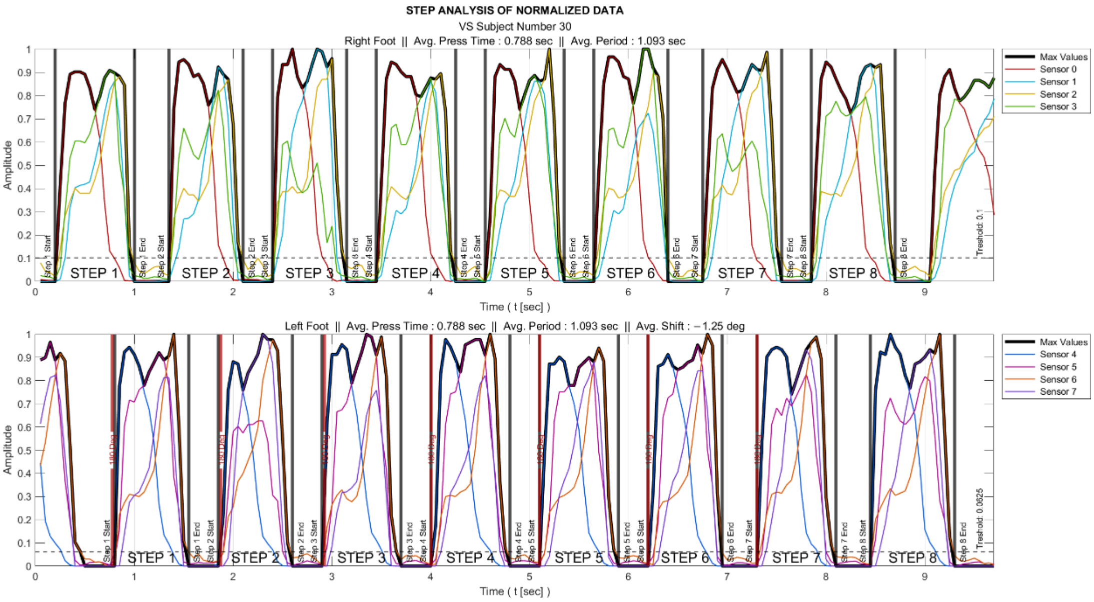


Figure 3.3: Segmentation of sensor data into sets of four rows for each foot, followed by combining these sets using the maximum operator and subsequently conducting step analysis.

3.2.3 Stage 3 – interpolation

At this stage, the interpolated normalized data undergoes individual processing for each sensor, eliminating segments that do not involve foot-ground contact. Subsequently, the data sequences are further refined by interpolating the sensor data to increase the resolution by a factor of 20. The rigorous process not only enhances the precision of data analysis but also facilitates obtaining a meaningful histogram representation with an adequate number of bins, each containing a sufficient number of samples, as depicted in Figure 3.4. The mentioned steps are demonstrated for a representative subject.

Linear interpolation was not preferred when adding new values to the dataset through interpolation in order to maintain accuracy without compromising the representation of the data. Instead, the cubic Hermite interpolation method was chosen as a preferred interpolation technique. This method provides a smoother and more accurate representation of the data while preserving its integrity [51].

Interpolation enhances the granularity of the data reveals subtle details that might be obscured at a lower resolution. Such an increase in resolution enables the detection of minute changes or deviations in gait patterns. By choosing the cubic Hermite interpolation over the linear method, we're prioritizing not just quantity but quality; the resulting data is a more genuine representation of the subjects' walking patterns.

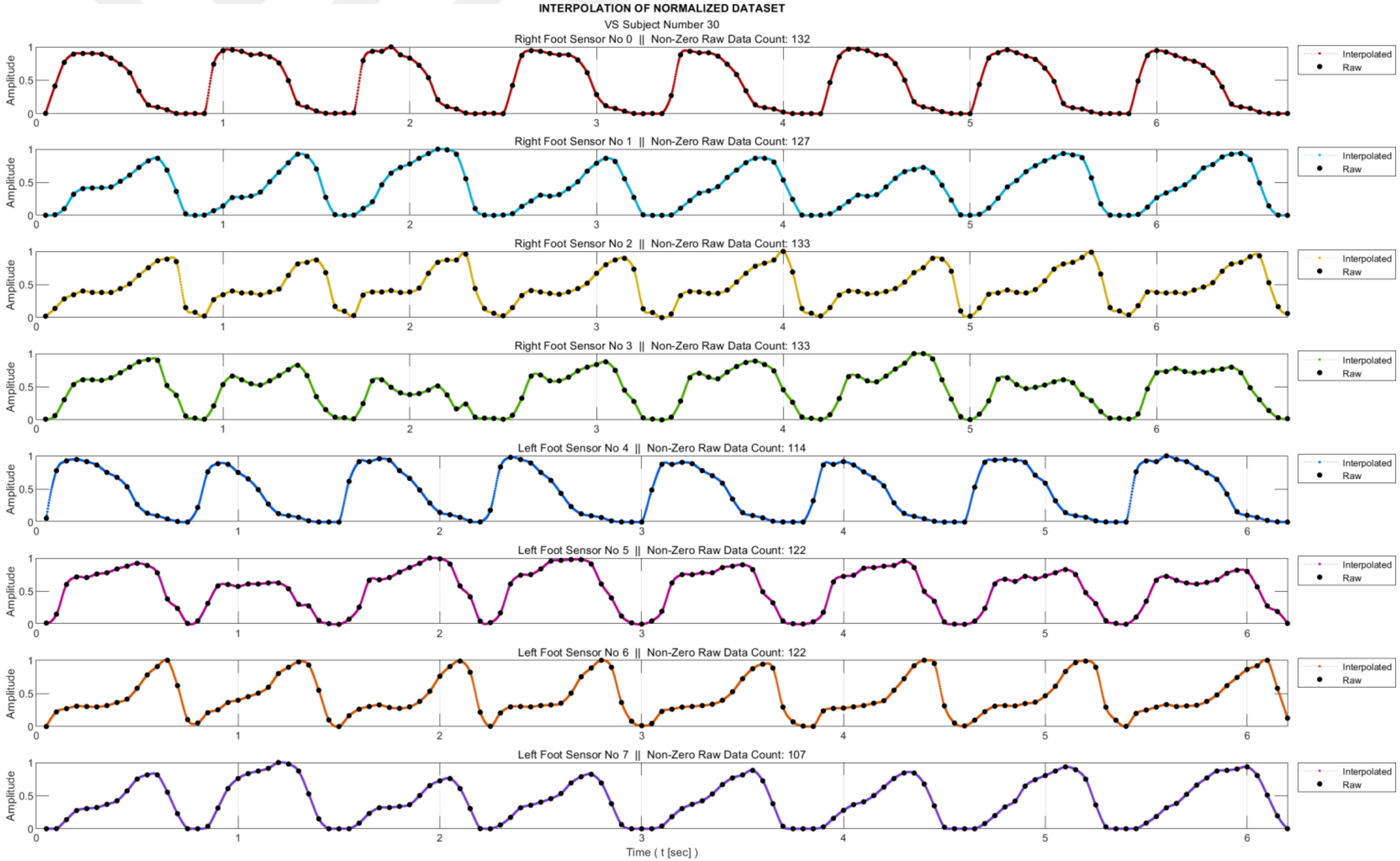


Figure 3.4: Removal of segments where the feet do not make contact with the ground from the data sequences, followed by interpolation applied to the data sequence.

3.2.4 Stage 4 – detrending

For each step, the trend of the interpolated data is calculated, and then the data is detrended based on these trend values.

Each sensor generates data during each step. However, the fact that the foot touches the ground does not mean that every sensor on the sole will generate data simultaneously. For instance, when data occur in the sensor located in the heel part of the insole, the sensor in the toe area may not yet generate any data. This phenomenon is clearly depicted in figure 3.4. Therefore, the starting and ending points of each step data are determined for each sensor data.

3.2.4.1 The proposed algoritm

In the study conducted, the aim was to extract features from various parameters. Therefore, the trend of the step data is derived by combining the previous trend curve with the step data itself. For the first step data in the interpolated data set, the trend curve is calculated as equivalent to itself since no trend has been identified before. This ensures that the entropy calculation of the detrended data is not affected by the first step. The process for the subsequent step data in the data array is as follows: The detrend curve of the previous step is scaled on the x-axis using the "Nearest-neighbor interpolation" method according to the specific step data. This process equalizes the data length of the previous trend curve and the length of the step data itself. Then, a trend curve is created for the specific step data by taking the weighted average of each data point with the alpha coefficient.

$$\begin{aligned} T_i &= F_i & i &= 1 \\ T_i &= \alpha F_i + (1-\alpha)\check{T}_{i-1} & i &= 1, 2 \dots n \end{aligned} \tag{3.4}$$

Here, T_i is the current-step trend data, and F_i stands for the current step data. $\check{T}$ denotes the trend data whose length is scaled, and α is a coefficient indicating the degree to which the previous trend curve is approximated to the current step data set. In Figure 3.5, α_{max} represents the maximum rate of change that each data point of the trend curve can exhibit from one step to the next, for which the value 0.23 was statistically determined, considering data from healthy subjects. We note that α_{max} serves as a parameter to achieve a balance between flexibility in trend curve adaptation and avoiding overfitting, and although it has a role in shaping the trend curve, the key features of our analysis remain relatively insensitive to its exact value. The process is terminated when the α value reaches α_{max} or the error value defined as $\varepsilon = mean\,\{|T_i - F_i|\}$ falls below a threshold so that it is considered negligible. The threshold level is set as 10^{-6}. The visual representation of the algorithm can be found in figure 3.5, and the outputs can be seen in figure 3.6.

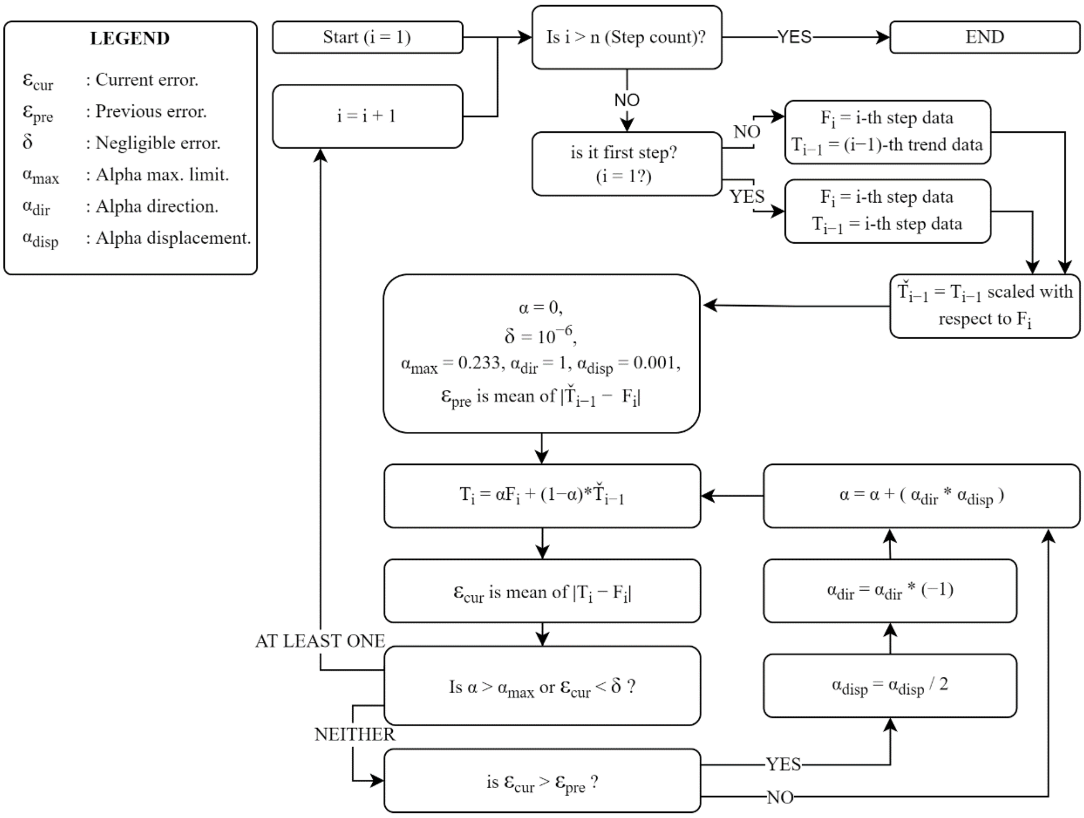
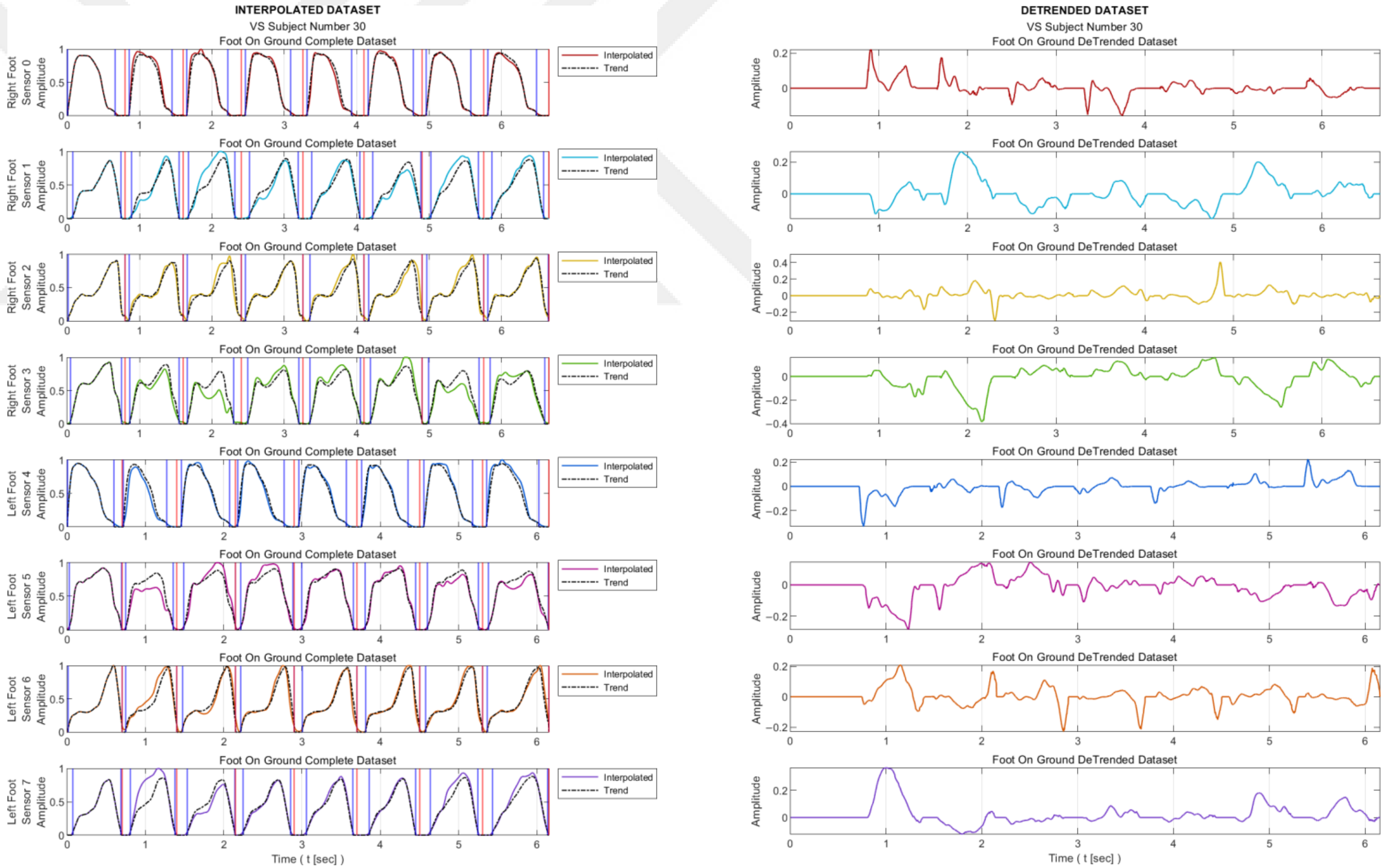


Figure 3.5: Flowchart of a specifically designed algorithm for generating the trend curve.

Detrending refines raw gait data into discernible patterns. Considering the unique activation timing of each sensor, we've devised an algorithm, showcased in figure 3.5, that blends past and present data trends for a holistic view. The heart of this method lies in the iterative adjustment of the weighted average coefficient (α) to minimize error. This adjustment starts from zero and tapers with each cycle until α peaks or the error reduction becomes trivial. With clear thresholds for α and error magnitude, our approach achieves a balance between precision and computational efficiency, ensuring the trend curve's adaptability and sensitivity to subtle gait changes. Consequently, we aim to capture the genuine intricacies of an individual's gait, filtering out any redundant data. The application of the resulting trend curves for detrending can be seen in figure 3.6.



(a) (b)

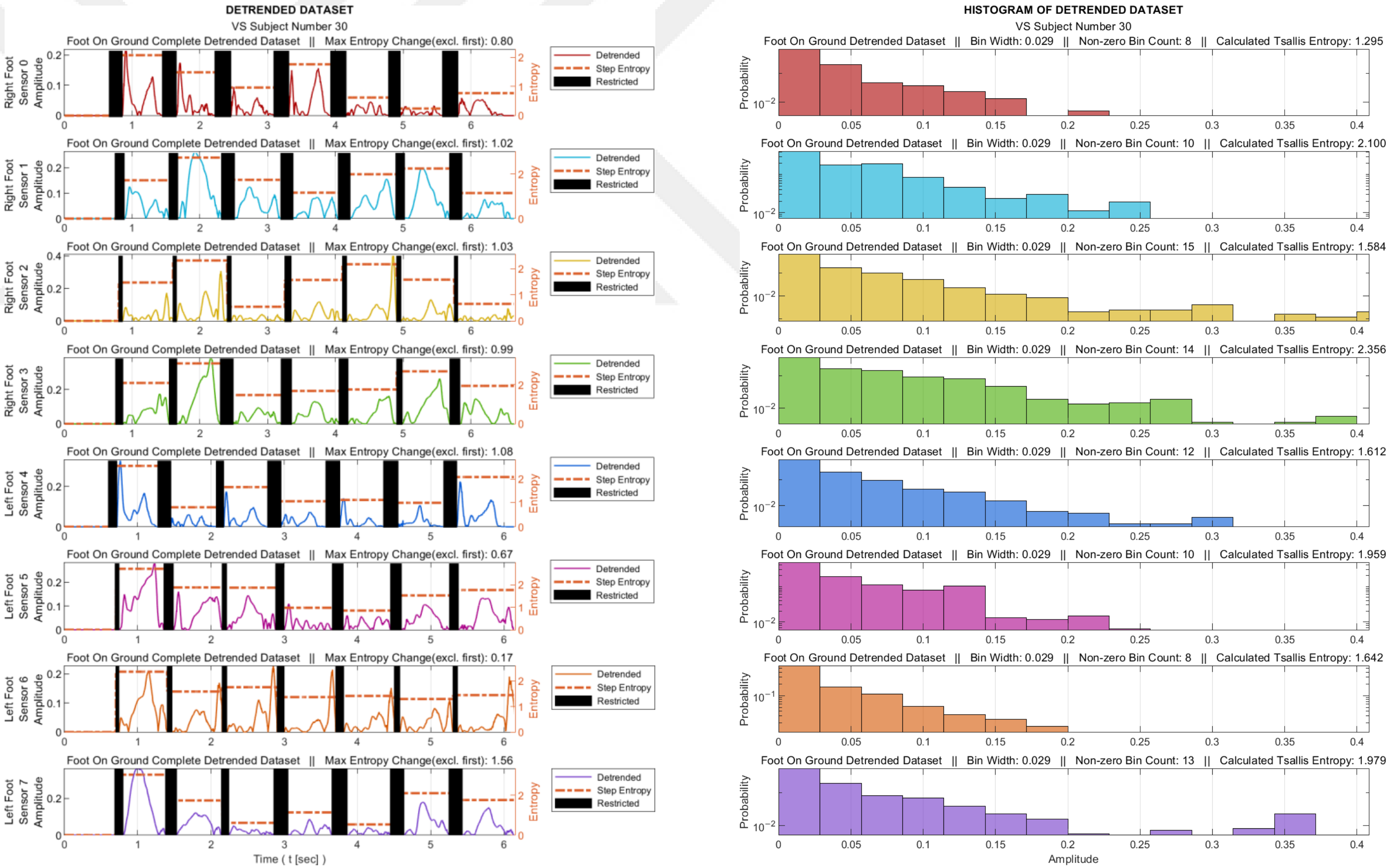
Figure 3.6: Detection of step data and calculation of **(a)** trend curves as well as **(b)** curves of detrended dataset for a sample VS-diseased subject. (Red vertical lines indicate the active stepping intervals of the foot; blue vertical lines indicate the active usage intervals of the relevant sensor).

3.2.5 Stage 5 – Tsallis entropy calculations

In this stage, the Tsallis entropy calculation was performed not only for the entire sensor data but also for each individual step within the entirety of the gait data. This approach allowed for a more detailed analysis of the step data and their relationship with the trend curve. By calculating the Tsallis entropy based on the detrended data, the study aimed to capture the information contained in the vertical displacements between the step data and the trend curve. The detrended data represented these distances, taking into account their magnitude and displacements. The absolute values of the detrended data were used to group data points with the same magnitude but different signs into the same histogram bin, enabling the calculation of probabilities and subsequent entropy measurements. Figure 3.7 provides a visual representation of the resulting histogram generated for all the detrended data taken from a VS diseased subject, demonstrating the distribution of data around the trend.

The choice of the number of bins in a histogram significantly impacts the level of detail and granularity in representing the data distribution. To achieve the desired resolution, a maximum of 25 bins was used in this study. Determining the maximum number of histogram bins was based on examining the detrended data of all subjects and identifying the maximum detrended size. By dividing the maximum detrended size by the maximum number of histogram bins, the precision of the range represented by each histogram bin was established. This approach ensured that the histogram effectively captured the variations in the detrended data and provided a comprehensive representation of the sensor data.

As explained in section 2.1, the value of the q parameter plays a crucial role in determining the nature of the entropy measurement. An empirical method was employed to determine the most suitable q value for feature extraction from the obtained datasets. Through extensive calculations and evaluations, a value of 0.82 was identified as the optimal q parameter. The success rates of the learning models were calculated for different Tsallis parameters (q values), using the specific classification models outlined in Stage 6 and employing a 10-fold cross-validation technique. The ratios of models achieving the highest success rates were used as the benchmark for determining the effectiveness of the Tsallis entropy method. These ratios are visually presented in figure 3.8, offering insights into the performance of different Tsallis parameters in classifying individuals with vestibular system disorder.



(a) (b)

Figure 3.7: For a sample diseased subject (no. 30): (a) Absolute values of the detrended data are illustrated in Figure 5b, and the step-by-step TE values are presented (Black bars indicate the segments where there is no data, representing them as a restricted area). The entropy values are calculated for each step; (b) Histograms derived from the entire gait data are plotted, with inactive sensor intervals removed.

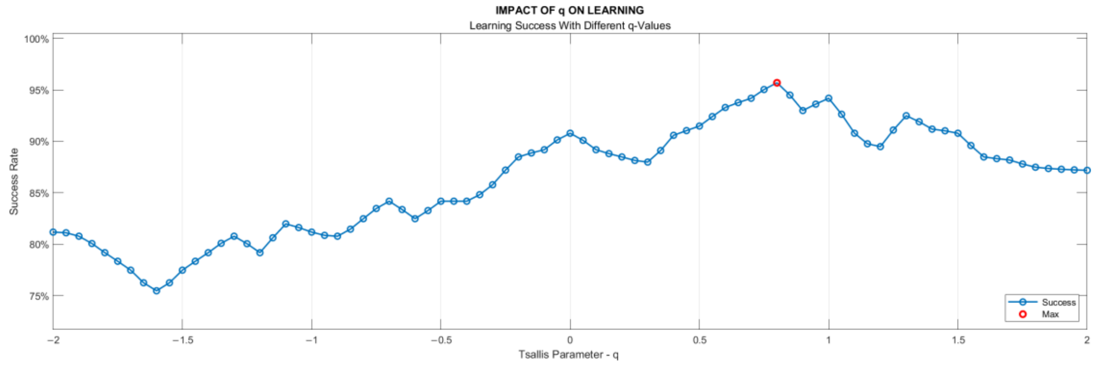


Figure 3.8: Correlation between learning success and varying Tsallis parameter (q) values.

Moreover, it is worth noting that the determination of the optimal q parameter value through empirical methods and subsequent evaluation of its impact on the performance of the learning models highlight the significance of parameter selection in entropy-based feature extraction.

At the completion of this stage, the study obtained Tsallis entropy values for all detrended data of each sensor, as well as the Tsallis entropy values calculated for each step within this detrended data. These entropy values provide quantitative measures of the complexity and information content of the gait patterns exhibited by individuals with vestibular system disorder. The incorporation of Tsallis entropy calculations and the examination of both sensor data and individual step data offer a comprehensive analysis of the gait dynamics, enabling a deeper understanding of the distinct characteristics of individuals with vestibular system disorder.

3.2.6 Stage 6 – feature extraction

The objective is to utilize the Tsallis entropy method on relatively short walking data in order to identify distinctive features that separate healthy individuals from those with diseases. The current stage involves establishing significant associations between the Tsallis entropy values of the overall gait for each sensor and the Tsallis entropy values calculated for the step data within each sensor data.

Obtaining a feature that indicates the distribution of entropy values from step to step is of great importance due to its capability to facilitate the identification and analysis of discrepancies, trends, and irregularities within a given data. However, considering that features were attempted to be extracted from relatively short gait and the number of steps were around ten in general, it was concluded that analysis methods yielding reliable results in a large number of data entries should be avoided. Therefore, the calculation of the deviation of the entropy values from a referance value from step to step was performed.

In the analysis of a flawless gait, the entropy levels and entropy change levels from step to step should be closer to zero compared to a flawed gait. For this reason, the deviation of the entropy values from step to step was calculated, with the goal of measuring how far they deviate from zero. The dataset containing the entropy values from step to step was augmented with the negation of all data values, and the standard deviation of the newly created dataset was calculated. This approach allows us to quantitatively measure the dispersion of the entropy values by calculating the standard deviation of the augmented dataset. It provides valuable insights into the extent to which the gait deviates from the desired zero entropy levels. The aforementioned procedures are expressed mathematically within equation 3.5.

$$E = \{e_1, e_2, \dots, e_n\} \; where \; e_k \in \mathbb{R} \; for \; k \in \mathbb{Z}^+$$

$$E' = \{e_1, e_2, \dots, e_n, -e_1, -e_2, \dots, -e_n\}$$

$$\sigma(E') = \sqrt{\frac{1}{2n} * \sum_{i=1}^{2n} (x_i - \mu)^2} \tag{3.5}$$

Here, e_k is Tsallis entropy value of k-th order step data. E denotes the ordered set of entropy values calculated step by step. E' symbolizes augmenting E with negative counterparts. $\sigma(E')$ represents the standard deviation of set E'. x_i stands for each element in set E'. μ signifies the mean of set E'. n is the total number of elements in set E.

3.2.7 Training of classifier models

In this final stage, machine learning was performed using sixteen features for each subject. These features comprised of eight total entropy levels for each sensor and eight deviations from zero values of entropies from step to step for each sensor.

The classifiers with the highest performances were determined using the Matlab R2021b Classification Learner Tool (on MSI GE75 Raider 10875H). The 10-fold cross validation technique was applied, where approximately 25% of the data (from 15 subjects) was used for testing, and the remaining data (from 45 subjects) was used for training.

The process of classification training involved utilizing nine different model categories provided by the Classification Learner Tool. These model categories utilized in this process include: Decision trees (DT), discriminant analysis, logistic regression, naïve bayes, support vector machine (SVM), nearest neighbors (KNN), kernel approximation, ensemble and neural networks. There are a total of thirty-two models in these model categories. For example, in the decision tree category, there are three models according to the maximum number of splits (Course: 4, Medium: 20, Fine: 100).

Among the all classifiers examined, SVM (with a Gaussian kernel), KNN(Cosine), and Logistic regression demonstrated superior performance. Brief explanations of these classifiers can be provided as follows: The KNN algorithm determines the class membership of an object/vector by examining its k nearest neighbors [52]. Logistic regression is a statistical model used to predict the probability of a dependent variable belonging to two or more classes in a dataset [53]. SVM seeks to find an optimal hyperplane to separate data clusters [54]. These three algorithms are commonly utilized in the literature when working with biomedical signals [55-60].

4. RESULTS AND CONCLUSIONS

The present study aimed to investigate the potential of Tsallis entropy calculations and feature extraction methods for diagnosing individuals with vestibular system disorders. The data processing methods employed in this study, including normalization, step analysis, interpolation, detrending, and Tsallis entropy calculations, were applied to obtain meaningful features from relatively short gait. By comparing the analysis data of VS diseased and healthy individuals, notable differences were observed, particularly in the feature extraction stage. The results indicated significant variations in entropy levels and step-wise entropy changes, suggesting the potential of these entropy-based features as objective indicators for the assessment of vestibular system disorders.

In the final stage of the study, machine learning techniques were utilized to classify the force data using the extracted features. SVM (Gaussian kernel), KNN (Cosine), and Logistic regression were identified as the classifiers that demonstrated the best performance, with success rates of 95%, 95%, and 93.3% respectively.

The findings of this study underscore the potential of Tsallis entropy calculations and feature extraction methods in diagnosing individuals with vestibular system (VS) disorders. This approach not only introduces a novel perspective for diagnosis and monitoring but also aims to enhance the accuracy and objectivity of assessments in clinical settings. Moreover, the results advocate for their integration in clinical scenarios for improved diagnosis, monitoring, and treatment evaluation. With further research and refinement, these techniques could lead to significant advancements in vestibular rehabilitation, ultimately elevating the quality of life for those with balance impairments.

4.1 Comprehensive Results Presentation

In this section, a comparative analysis will be conducted between the data collected from individuals with a disorder in the Vestibular system and that obtained from healthy subjects. The comparison will commence from the fourth stage of a sensor data, as described in the section 3.2 section. This comparative approach is integral to understanding the nuances and intricacies of the sensory data and how it reflects the physiological disparities between the two groups.

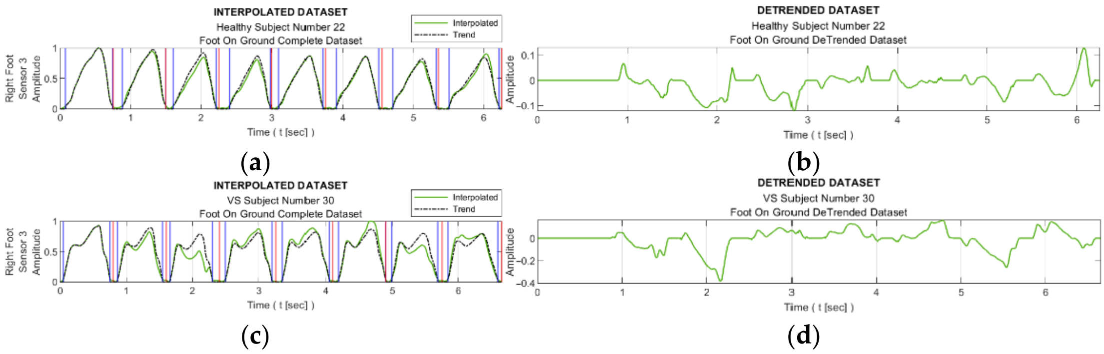


Figure 4.1: (a, c) Visualization of interpolated sensory data and the trend curve of healthy and VS diseased subjects. **(b)** Visualization of detrended data from a healthy subject. **(d)** Visualization of detrended data from a subject with a vestibular system disorder.

Figure 4.1 facilitates the observation of discernible variations in the behavior of sensor S3 when comparing healthy and VS diseased individuals. Additionally, it visually presents the detrended data, which represents the disparity between step data and the trend curve.

Furthermore, figure 4.1 provides a comprehensive visual analysis of sensor S3, enabling the identification of significant differences between healthy and VS diseased individuals. The graphical representation depicted in figure 4.2 illustrates a notable distinction in the bin distributions of histogram between a healthy individual and an individual diagnosed with a vestibular system disorder.

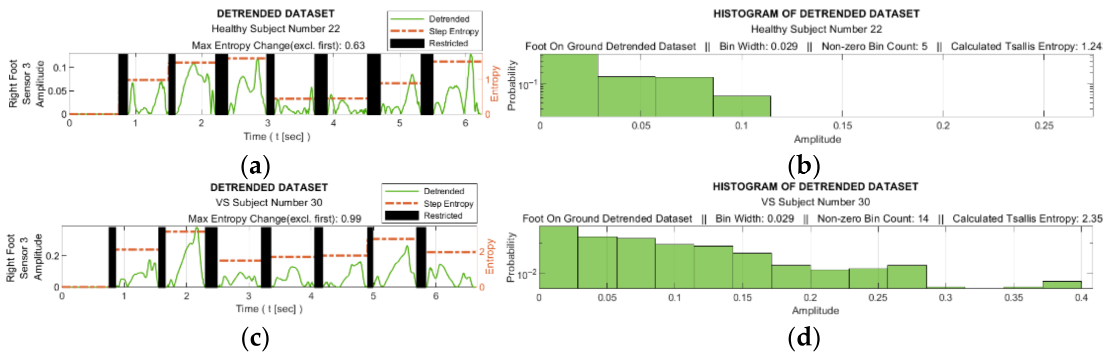


Figure 4.2: (a) and **(c)** Visualization of absolute expressions of the detrend data, restricting of the data-free segments from the dataset, and step-by-step entropy values of a healthy subject and VS diseased subject. **(b)** and **(d)** Visualization of a histogram depicting the absolute expression of detrended data obtained from a healthy subject and VS diseased subject.

These visual representations provided by figure 4.1 and figure 4.2 enable a comprehensive analysis of gait patterns and sensor behavior in individuals with vestibular system disorders compared to healthy individuals. By examining the discernible variations in sensor S3's behavior in figure 4.1, one can gain insights into the impact of vestibular system disorders on locomotion. Moreover, figure 4.2 further highlights the significant disparity in bin distributions of the histogram, emphasizing the distinct characteristics between a healthy individual and someone diagnosed with a vestibular system disorder. Such visual analyses are pivotal as they translate intricate numerical data into more digestible visual formats. This facilitates quicker decision-making processes for clinicians and provides researchers with a clearer direction for subsequent investigations. Together, these graphical representations offer valuable visual cues for understanding and comparing the effects of vestibular system disorders on gait and sensor data.

In Figure 4.2, the detrended data graphs, derived from figures 4.1b and 4.1d, are presented with their absolute values. The histograms generated from these graphs can also be observed. Notably, in Figures 4.2a and 4.2c, the black bars signify the inactive durations associated with the respective sensor.

Based on the selected sample subjects and their corresponding sensor data from the research study, the maximum incremental change in the TE value was observed to be 0.63 for a healthy participant. Conversely, a value of 0.99 was recorded for an individual diagnosed with a vestibular system disorder. For the entire gait cycle, the comprehensive TE value was determined to be 1.243 for the healthy individual, while a value of 2.356 was registered for the affected party. All TE values for these sample subjects, based on the sensor data, are tabulated in Table 4.1. A consolidated visualization of the TE values for the entire gait cycle across all participants is presented in figure 4.3.

Table 4.1: TE values calculated from each sensor's data for sample subjects.

	Healthy Subject (no. 22)		VS Subject (no. 30)	
Sensor	**Entire Gait**	**Stepwise Max**	**Entire Gait**	**Stepwise Max**
S0	1.39	0.98	1.29	0.80
S1	2.15	0.83	2.10	1.02
S2	1.38	0.72	1.58	1.03
S3	1.24	0.63	2.36	0.99
S4	1.08	0.87	1.61	1.08
S5	1.38	0.79	1.96	0.67
S6	1.36	0.82	1.64	0.17
S7	1.54	0.86	1.98	1.56

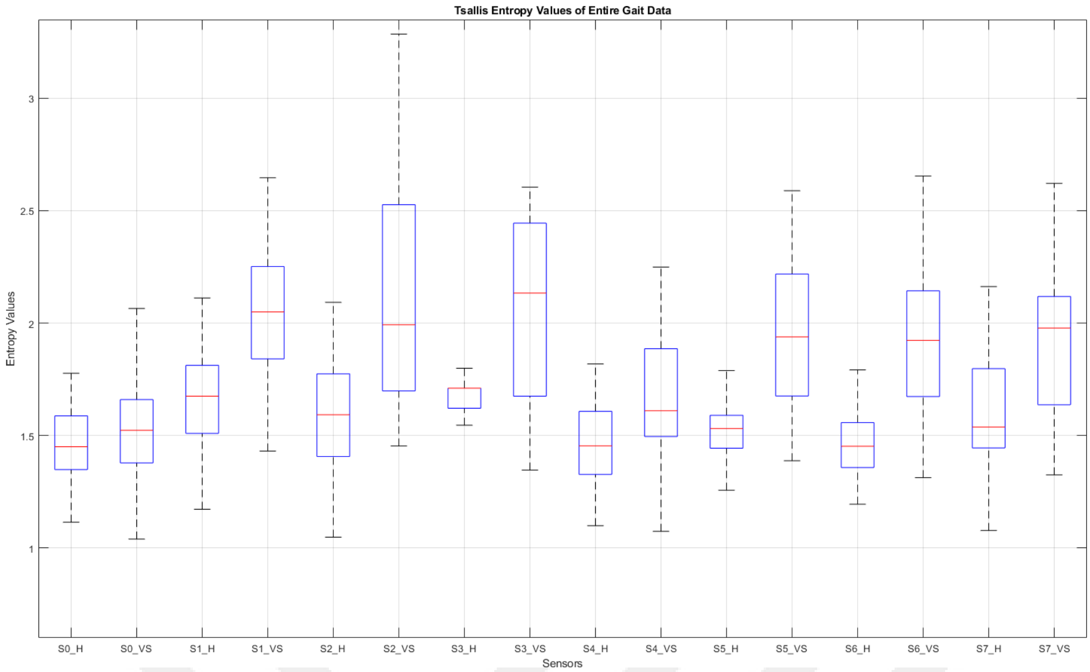


Figure 4.3: Box plot of the entire-gait TE values for all participants. S: sensor, H: healthy, VS: diseased.

As delineated in section 3.2.7, thirty-two classifiers from Matlab's Classification Learner Tool were trained using sixteen features per subject. This training process employed sixteen distinctive features for each individual subject. Notably, these features were constituted by eight distinct total entropy levels, one for each sensor. Additionally, there were eight deviations from the zero values of entropies, which were measured from one step to the next for every sensor in the set. The average accuracies of the principal classification algorithms are catalogued in table 4.2. For deeper insights, table 4.3 showcases the confusion matrices, while figure 4.4 offers a visual representation through the corresponding Receiver Operating Characteristic (ROC) curves. These are particularly for one of the ten training-test set pairs that represent the top-performing three classifiers.

Table 4.2: Accuracy of Major Classification Algorithms

Algorithm	Accuracy (%)
SVM (Gaussian)	95.0
Logistic Regression	95.0
KNN (Cosine)	93.3
Neural Network (Wide)	93.3
Kernel (SVM)	91.7
Ensemble (Bagged Tree)	88.3
Naïve Bayes (Kernel)	86.7
Quadratic Discriminant	78.3
Decision Tree (Fine)	73.3

Table 4.3: Confusion Matrices for One Of The Ten Training-Test Set Pairs

Predicted Class	SVM (Gaussian)		Logistic Regression		KNN (Cosine)	
	H	D	H	D	H	D
H	30	0	29	1	28	2
D	3	27	2	27	2	28

H – Healthy, D – Diseased

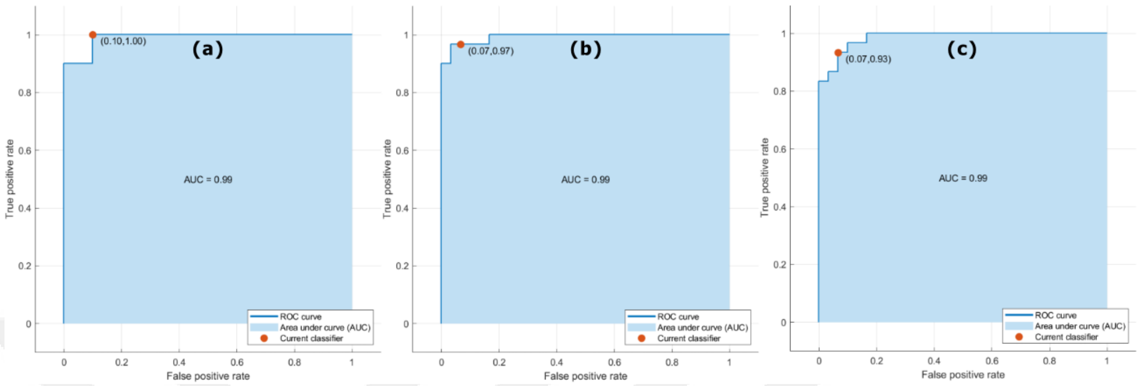


Figure 4.4: ROC curves associated with **(a)** the support vector machine (SVM) model with Gaussian kernel, **(b)** logistic regression, and **(c)** the k-nearest neighbors (KNN) algorithm using cosine similarity in Table 4.3.

Table 4.4 offers an intricate exploration into the efficacy and reliability of the top two classification algorithms. This in-depth tabulation serves as an instrumental tool in quantifying the performance metrics of these algorithms. The need for such a comprehensive table stems from earlier observations that highlighted distinct variations in sensor behaviors between healthy individuals and those diagnosed with vestibular system disorders. Such variations underscore the intricacies and complexities of the underlying data patterns. Given this context, it becomes imperative to not just rely on surface-level observations but to delve deeper into performance metrics. Moreover, in the evolving realm of data science and medical diagnosis, leveraging advanced algorithms is key to ensuring accurate and timely patient assessment. This ensures that the classification algorithms employed are not just theoretically apt, but they also demonstrate empirical effectiveness in differentiating between the two distinct groups. By scrutinizing these metrics, researchers and practitioners can achieve a heightened confidence in the algorithms' capacity to discern and categorize data, thereby solidifying the scientific robustness of the study.

The metrics highlighted in Table 4.4 not only validate the robustness and precision of the SVM (Gaussian) and Logistic Regression algorithms, but they also emphasize their ability to effectively distinguish between the unique sensor patterns of the two studied groups.

Table 4.4: Some statistical data about the top two classification algorithms.

Statistical Property	SVM (Gaussian)	Logistic Regression
accuracy (%)	95.0	95.0
sensitivity (%)	91.6	94.0
specificity (%)	97.9	95.1
F1 Score	0.945	0.943
MCC	0.899	0.891

4.2 Performance Evaluation of the Proposed Algorithm

The comparison of various trend generation algorithms in machine learning is paramount to understanding the efficacy of the algorithms being developed. For the purpose of this study, and to ascertain the effect of the proposed trending algorithm, trend curves were meticulously created using 2nd, 3rd, and 4th-degree curve-fitting polynomials. Subsequently, their influence on the performance and success rates of machine learning models was critically evaluated. The success percentages, along with classification accuracies achieved using these distinct trend generation methods, are comprehensively presented in Table 4.4.

Table 4.5: Classification accuracies with different trend generation methods.

Classification Model	Proposed Algorithm	Fourth Degree Polynomial	Third Degree Polynomial	Second Degree Polynomial
SVM – Gaussian	95.0%	81.7%	76.3%	71.7%
Logistic Regression(LR)	95.0%	76.3%	78.3%	63.3%
KNN – Cosine	93.3%	78.3%	70.0%	66.7%
Best Method	95.0% (with SVM-G & LR)	86.7% (with Ensemble Subsp. Discr.)	83.3% (with Decision Trees-Fine/Medium)	83.3% (with Ensemble-Bagged Trees)

The results of this comparison highlight the significance of the proposed trend generation algorithm. By developing an algorithm specifically tailored for trend generation in the context of the analyzed data, it is possible to achieve higher success percentages in machine learning tasks compared to using traditional polynomial-based approaches. This demonstrates the effectiveness of the developed algorithm in capturing the underlying trends and patterns present in the data.

The ability to accurately capture trends is crucial in machine learning as it directly impacts the performance and predictive capabilities of the models. By employing the proposed algorithm, which is designed to generate trend curves optimized for the characteristics of the dataset, the machine learning models can better capture and utilize the inherent trends in the data, leading to improved accuracy.

Overall, the proposed trend generation algorithm plays a crucial role in enhancing the performance of machine learning models by accurately capturing the underlying trends in the data. Its effectiveness, adaptability, and ability to outperform traditional polynomial-based approaches highlight its importance and its potential for improving various applications, ranging from data analysis to predictive modeling.

4.3 Visual Comparison of Gait Processing

In this section, a comparative analysis is presented, based on force data derived from two groups: healthy individuals and those diagnosed with vestibular system disorders. The analysis leverages data processing techniques and feature extraction methods to reveal distinct differences in gait patterns between these groups. These findings are significant, as they demonstrate the effectiveness of these methods in identifying variations in gait, which is crucial for diagnosing and monitoring vestibular system disorders. This analysis offers healthcare professionals a valuable tool for the assessment and management of such conditions.

To provide a comprehensive understanding, the section includes a series of figures in Appendix A, labelled from Figure A.1 to A.7. Each figure visually represents the differences in gait patterns between the two groups, illustrating the impact of each data processing step. This visual representation aids in highlighting the nuanced differences in gait patterns, further supporting the findings of the comparative analysis.

4.4 Summary of Findings

The findings of this study demonstrate the effectiveness of Tsallis entropy calculations and feature extraction methods in analysis of gait data for individuals with vestibular system disorders. The gait data was collected using insole force sensors, offering a detailed and reliable source of information. The data collection was optimized to a short span of 10-15 seconds, which is considerably shorter than most experiments in the literature. The data processing pipeline successfully generated meaningful features that captured distinctive gait patterns in individuals with vestibular system disorders. Through machine learning classification, the study achieved high accuracy in differentiating between healthy individuals and those with vestibular system disorders. This highlights the potential of entropy-based features as objective indicators for the diagnosis and monitoring of vestibular system disorders.

4.5 Final Conclusion and Key Insights

In conclusion, this study provides evidence supporting the use of Tsallis entropy calculations and feature extraction methods as potential tools for analysis of gait data in individuals with vestibular system disorder. The comprehensive data processing approach employed in this research facilitated the identification of significant differences in pressure analysis data between healthy individuals and those with vestibular system disorder. The extracted entropy-based features demonstrated their potential as objective indicators for the assessment of vestibular system disorders.

Moreover, the classification results obtained through machine learning techniques underscored the discriminative power of these features. SVM (Gaussian kernel), KNN (Cosine), and Logistic regression emerged as the top-performing classifiers, further supporting the utility of entropy-based features in distinguishing between healthy individuals and those with vestibular system disorder.

These findings hold implications for clinical practice, suggesting the integration of Tsallis entropy calculations, machine learning algorithms, and longitudinal monitoring in the diagnosis and monitoring of vestibular system disorders. By adopting these recommendations, clinicians can enhance their understanding of gait patterns, improve diagnostic accuracy, and make informed treatment decisions.

Employing Tsallis entropy in gait data analysis reveals intricate nuances of gait patterns specific to individuals with vestibular system disorders. Furthermore, the integration of machine learning, especially classifiers like SVM, KNN, and Logistic regression, adds substantial value; these tools adeptly discern subtle differences in gait patterns, presenting a notable advancement in diagnostic methodologies. By synergizing Tsallis entropy with machine learning, this research proposes a more objective and data-driven approach in diagnosing and monitoring vestibular system disorders.

Further research involving larger-scale studies and diverse populations is warranted to validate the findings and explore the applicability of these methods in different disease conditions. Continued advancements in gait data analysis techniques offer promising avenues for improving the assessment, diagnosis, and management of vestibular system disorders, ultimately leading to better patient outcomes.

5. DISCUSSION AND RECOMMENDATIONS

This thesis presents the key findings and implications of a study investigating the potential applications of Tsallis entropy calculations and feature extraction methods in the diagnosis and monitoring of individuals with vestibular system disorders. The study employed a comprehensive data processing pipeline to analyze gait data and extract meaningful features for characterizing gait patterns. Machine learning techniques were utilized to classify individuals as healthy or having a vestibular system disorder based on the extracted features. The results highlight the significance of entropy-based features in distinguishing between the two groups. This conclusion section summarizes the main findings, discusses the implications of the study, explores potential applications in vestibular system disorder diagnosis, addresses limitations, and suggests future research directions.

5.1 Discussion on Entropy-Based Findings

This study aims to discover novel features for the identification of diseases related to the vestibular system, with the objective of enhancing detection accuracy and reducing the duration of data acquisition. In this direction, the objective is to utilize the Tsallis entropy method on relatively short walking data to extract features that can differentiate between healthy and diseased individuals. Based on the results of this study, noticeable differences were shown in gait data. Feature extraction using Tsallis entropy calculations obtained through data processing methods has emerged as a potential tool for the diagnosis and monitoring of individuals with vestibular system disorder.

The data processing methods used in this study aim to obtain meaningful features from relatively short gait. Six stages, namely normalization, step analysis, interpolation, detrending, Tsallis entropy calculations, and feature extraction, were employed to process the data and perform a detailed analysis. These stages were applied to present the dataset in a more understandable and concise manner. These features play a significant role in identifying the characteristics of gait patterns in individuals with vestibular system disorder during pressure analysis.

Tsallis entropy calculations were performed based on the histograms of detrended data. The aim was to establish meaningful connections using the total entropy levels for each sensor and the deviations of step-wise entropy changes from zero. The measurement of how much the step-wise deviations of entropy values deviate from zero was used for feature extraction. The results indicate notable differences, particularly in the feature extraction stage where Tsallis entropy calculations and step-wise entropy changes were evaluated. Compared to healthy individuals, the entropy levels and step-wise entropy changes in pressure analysis data of individuals with vestibular system disorder show significant variations. These findings highlight the importance of entropy-based features that can be used as potential indicators for the objective assessment of vestibular system disorders.

In the final stage, machine learning was performed using sixteen features for each subject through classification learning. The aim was to identify classifiers capable of distinguishing between healthy individuals and those with vestibular system disorder. SVM (Gaussian kernel), KNN (Cosine), and Logistic regression were identified as classifiers that showed the best performance.

Based on the results of this study, Tsallis entropy calculations and feature extraction methods appear to be potential tools for analysis of gait data in individuals diagnosed with vestibular system disorder. This approach, used to detect differences between healthy individuals and those with vestibular system disorder, may offer a new perspective for the diagnosis and monitoring of vestibular system disorders.

5.2 Implications of the Study

The implications of this study are twofold. Firstly, the utilization of Tsallis entropy calculations and feature extraction methods provides a new approach for analyzing gait data in individuals with vestibular system disorders. This objective assessment tool can assist healthcare professionals in diagnosing and monitoring the condition, leading to improved treatment and management strategies. Secondly, the study highlights the importance of entropy-based features in gait data analysis research, not only for vestibular system disorders but also for other related neurological conditions.

The findings open doors for further exploration and application of entropy-based analysis in various clinical settings.

5.3 Limitations

While this study provides valuable insights into the potential of Tsallis entropy calculations and feature extraction methods, there are limitations to consider. Firstly, the study focused on a specific group of individuals with vestibular system disorders, limiting the generalizability of the findings to broader populations. Future research should include larger sample sizes and consider individuals with various types and severities of vestibular system disorders. Additionally, the study did not explore the influence of factors such as stress or comorbidities on results, which could be potential areas for further investigation.

Furthermore, the study primarily compared data from VS diseased individuals to healthy individuals. Future studies should explore the differentiation of specific vestibular disorders and compare the results to relevant clinical measures. Moreover, the study solely focused on analysis of force data; future research could investigate the potential of Tsallis entropy calculations in other aspects of vestibular system assessment, such as balance and posture analysis. Nonetheless, the findings of this study provide valuable insights into the potential utility of Tsallis entropy calculations and feature extraction methods for diagnosis VS diseased individuals, opening up new possibilities for enhanced diagnosis and treatment strategies in this population.

Further research with larger-scale studies and a broader range of disease conditions is needed to address these limitations and enhance the validity and applicability of the proposed methods.

5.4 Future Study

The current study provides a foundation for future research in the field of gait data analysis for vestibular system disorders. Further investigations should focus on validating the effectiveness of Tsallis entropy calculations and feature extraction methods using larger and more diverse participant groups. Additionally, the application of these methods to different disease conditions within the realm of vestibular dysfunction would enhance the understanding of the underlying gait abnormalities. Longitudinal studies tracking the progression of vestibular system disorders and the evaluation of treatment outcomes using these methods could provide valuable insights into disease management and rehabilitation strategies.

5.5 Potential Applications in Vestibular System Disorder Diagnosis

In this study, an innovative algorithm was developed to determine the trend curve at each step, which showed superior performance in comparison to other curve fitting methods. This can enhance the analysis and accuracy of gait patterns in clinical practice.

The potential applications of Tsallis entropy calculations and feature extraction methods in vestibular system disorder diagnosis are significant. By incorporating these methods into clinical practice, healthcare professionals can obtain quantitative and objective measurements of gait patterns. This can aid in the accurate diagnosis of vestibular system disorders, allowing for early intervention and personalized treatment plans. Additionally, the objective nature of entropy-based features reduces subjectivity in the diagnosis process and enhances the overall reliability of assessments.

5.6 Recommendations for Clinical Practice

The findings of this study have significant implications for clinical practice in the diagnosis and monitoring of individuals with vestibular system disorder. The utilization of Tsallis entropy calculations and feature extraction methods can provide valuable insights into gait patterns and serve as objective assessment tools. These techniques can aid clinicians in the accurate diagnosis and ongoing monitoring of vestibular system disorders. Based on the results, we recommend the following practices for clinical application:

1. Incorporating Tsallis entropy calculations: Clinicians should consider integrating Tsallis entropy calculations into their assessment protocols for gait data analysis. This approach can provide additional quantitative measures to complement subjective evaluations, enabling a more comprehensive understanding of gait patterns.
2. Integration of machine learning algorithms: The utilization of machine learning algorithms, such as SVM (Gaussian kernel), KNN (Cosine), and Logistic

regression, can enhance the diagnostic accuracy and efficiency of analysis of gait data in individuals with vestibular system disorder. Clinicians should familiarize themselves with these algorithms and consider their implementation in clinical settings.

3. Longitudinal monitoring: Given the potential of entropy-based features in tracking changes in gait patterns, longitudinal monitoring of individuals with vestibular system disorder is recommended. Regular assessments using Tsallis entropy calculations and feature extraction methods can help evaluate the effectiveness of interventions, track disease progression, and inform treatment decisions.

REFERENCES

[1] **Khan, S. and Chang, R.** (2013). Anatomy of the vestibular system: A review. NeuroRehabilitation. 32, 437-443. doi:10.3233/NRE-130866.

[2] **Tao, W.; Liu, T.; Zheng, R. and Feng, H.** Gait analysis using wearable sensors. Sensors (Basel). 2012;12(2):2255-83. doi: 10.3390/s120202255. Epub 2012 Feb 16. PMID: 22438763; PMCID: PMC3304165.

[3] **Dougherty, J.M.; Carney, M.; Hohman, M.H. and et al.** Vestibular Dysfunction. [Updated 2023 Mar 1]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK558926/

[4] **Ghai, S.; Hakim, M.; Dannenbaum, E. and Lamontagne, A.** Prevalence of Vestibular Dysfunction in Children with Neurological Disabilities: A Systematic Review. Front Neurol. 2019 Dec 17; 10:1294. doi: 10.3389/fneur.2019.01294. PMID: 31920918; PMCID: PMC6928113.

[5] **Wagner, A.R.; Akinsola, O.; Chaudhari, W. and et al.** Measuring Vestibular Contributions to Age-Related Balance Impairment: A Review, Frontiers in Neurology, 12, 2021, DOI=10.3389/fneur.2021.635305, ISSN=1664-2295.

[6] **Hall, C.D.; Herdman, S.J.; Whitney, S.L.; Cass, S.P. and et al.** Vestibular Rehabilitation for Peripheral Vestibular Hypofunction: An Evidence-Based Clinical Practice Guideline: From The American Physical Therapy Association Neurology Section. J Neurol Phys Ther. 2016 Apr;40(2):124-55. doi: 10.1097/NPT.0000000000000120. PMID: 26913496; PMCID: PMC4795094.

[7] **Marsden, J.; Pavlou, M.; Dennett, R. and et al.** Vestibular rehabilitation in multiple sclerosis: study protocol for a randomised controlled trial and cost-effectiveness analysis comparing customised with booklet based vestibular rehabilitation for vestibulopathy and a 12 month observational cohort study of the symptom reduction and recurrence rate following treatment for benign paroxysmal positional vertigo. BMC Neurol 20, 430 (2020). https://doi.org/10.1186/s12883-020-01983-y

[8] **Vanicek, N.; King, S.A.; Gohil, R.; Chetter, I.C. and Coughlin, P.A.** Computerized dynamic posturography for postural control assessment in patients with intermittent claudication. J Vis Exp. 2013 Dec 11;(82): e51077. doi: 10.3791/51077. PMID: 24378378; PMCID: PMC4047968.

[9] **Buckley, C.; Alcock, L.; McArdle, R.; Rehman, R.Z.U. and et. al.** The Role of Movement Analysis in Diagnosing and Monitoring Neurodegenerative Conditions: Insights from Gait and Postural Control. Brain Sci. 2019

Feb 6;9(2):34. doi: 10.3390/brainsci9020034. PMID: 30736374; PMCID: PMC6406749.

[10] **Higuma M.; Sanjo, N.; Mitoma, H.; Yoneyama, M. and Yokota**, **T.** “Wholeday gait monitoring in patients with Alzheimer’s disease: A relationship between attention and gait cycle,” J. Alzheimer’s Dis. Rep., vol. 1, no. 1, pp. 1–8, 2017.

[11] **Nieto-Hidalgo M.; Ferrández-Pastor, F. J. and et al,** “Gait analysis using computer vision based on cloud platform and mobile device,” Mobile Inf. Syst., vol. 2018, pp. 1–10, Jan. 2018, Art. no. 7381264.

[12] **Cicirelli, G.; Impedovo, D.; Dentamaro, V.; Marani, R. and et al.** "Human Gait Analysis in Neurodegenerative Diseases: A Review," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 1, pp. 229-242, Jan. 2022, doi: 10.1109/JBHI.2021.3092875.

[13] **Giladi, N.; Horak, F.B. and Hausdorff, J.M.** Classification of gait disturbances: distinguishing between continuous and episodic changes. Mov Disord. 2013 Sep 15;28(11):1469-73. doi: 10.1002/mds.25672. PMID: 24132835; PMCID: PMC3859887.

[14] **Bovonsunthonchai, S.; Vachalathiti, R. and Hiengkaew, V.** et al. Quantitative gait analysis in mild cognitive impairment, dementia, and cognitively intact individuals: a cross-sectional case–control study. BMC Geriatr 22, 767 (2022). https://doi.org/10.1186/s12877-022-03405-9

[15] **Guo, Y.; Yang, J. and Liu, Y., et al.** Detection and assessment of Parkinson's disease based on gait analysis, Frontiers in Aging Neuroscience, 14, 2022, doi: 10.3389/fnagi.2022.916971, ISSN=1663-4365

[16] **Ikizoglu, S. and Heydarov, S.** Accuracy comparison of dimensionality reduction techniques to determine significant features from IMU sensor-based data to diagnose vestibular system disorders. Biomedical Signal Processing and Control. 2020, 61. 101963. 10.1016/j.bspc.2020.101963.

[17] **Agrawal D.K.; Usaha, W.; Pojprapai, S. and Wattanapan, P.** "Fall Risk Prediction Using Wireless Sensor Insoles with Machine Learning," in IEEE Access, vol. 11, pp. 23119-23126, 2023, doi: 10.1109/ACCESS.2023.3252886.

[18] **Bustamante, P.; Solas, G. and Grandez, K.** Neurodegenerative Disease Monitoring Using a Portable Wireless Sensor Device. InTech. 2011 doi: 10.5772/29814.

[19] **Schniepp, R.; Möhwald, K. and Wuehr, M.** Clinical and automated gait analysis in patients with vestibular, cerebellar, and functional gait disorders: perspectives and limitations. J Neurol. 2019 Sep;266(Suppl 1):118-122. doi: 10.1007/s00415-019-09378-x. Epub 2019 May 27. PMID: 31134375.

[20] **Mohan, D.M.; Khandoker, A.H.; Wasti, S.A. and et al.** Assessment Methods of Post-stroke Gait: A Scoping Review of Technology-Driven Approaches to Gait Characterization and Analysis. Front Neurol. 2021

Jun 8;12:650024. doi: 10.3389/fneur.2021.650024. PMID: 34168608; PMCID: PMC8217618.

[21] **Das, R.; Paul, S.; Mourya, G.K.; Kumar, N. and Hussain, M.** Recent Trends and Practices Toward Assessment and Rehabilitation of Neurodegenerative Disorders: Insights From Human Gait. Front Neurosci. 2022 Apr 15;16:859298. doi: 10.3389/fnins.2022.859298. PMID: 35495059; PMCID: PMC9051393.

[22] **Wagner, A.R. and Reschke, M.F.** Aging, vestibular function, and balance control: physiological and behavioral considerations. Current Opinion in Physiology, 2021, 19, 67-74, doi=10.3389/fneur.2021.635305, ISSN=1664-2295

[23] **Schmidheiny, A.; Swanenburg, J.; Straumann, D. et al.** Discriminant validity and test re-test reproducibility of a gait assessment in patients with vestibular dysfunction. BMC Ear Nose Throat Disord 15, 6 (2015). https://doi.org/10.1186/s12901-015-0019-8

[24] **Tsallis, C.** Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World, Contemporary Physics, pp. 431-438, 2009. DOI: 10.1080/00107510902868668

[25] **Al-Nuaimi, A.H.; Jammeh, E.; Sun, L. and Ifeachor, E.** Tsallis entropy as a biomarker for detection of Alzheimer's disease. Annu Int Conf IEEE Eng Med Biol Soc. 2015; 2015:4166-9. doi: 10.1109/EMBC.2015.7319312. PMID: 26737212.

[26] **Gao, J.; Hu, J. and Tung, W.** Entropy measures for biological signal analyses. Nonlinear Dynamics. 2012, 68. 10.1007/s11071-011-0281-2.

[27] **Liu, A.B. and Lin, C.W.** Multiscale Approximate Entropy for Gait Analysis in Patients with Neurodegenerative Diseases. Entropy. 2019; 21(10):934. https://doi.org/10.3390/e21100934

[28] **Haid, T. and Federolf, P.** Human Postural Control: Assessment of Two Alternative Interpretations of Center of Pressure Sample Entropy through a Principal Component Factorization of Whole-Body Kinematics. Entropy. 2018; 20(1):30. https://doi.org/10.3390/e20010030

[29] **Davalos-Bichara, M.; Zuniga, M.G.; Agrawal, Y. and et al.** Forward and backward locomotion in individuals with dizziness Gait & Posture, Volume 40, Issue 4, 2014, Pages 499-503, ISSN 0966-6362, https://doi.org/10.1016/j.gaitpost.2014.06.008.

[30] **Schwämmle, V. and Tsallis, C.** Two-parameter generalization of the logarithm and exponential functions and Boltz-mann-Gibbs-Shannon entropy. J. Math. Phys. 2007; 48(11). DOI: 10.1063/1.2801996

[31] **Liang, Z.; Wang, Y.; Sun, X. and et al. (2015)** Entropy Measures in Anesthesia, Frontiers in Computational Neuroscience, vol. 9, 16. DOI=10.3389/fncom.2015.00016, ISSN=1662-5188.

[32] **Xiong, W.; Faes, L. and Ivanov, P.C.** Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of

artifacts, nonstationarity, and long-range correlations. Phys Rev E. 2017 Jun;95(6-1):062114. doi: 10.1103/PhysRevE.95.062114.

[33] **Li, C. and Shang, P.** Multiscale Tsallis permutation entropy analysis for complex physiological time series, Physica A: Statistical Mechanics and its Applications. 2019, vol 529, Pages 10-20, ISSN 0378-4371, https://doi.org/10.1016/j.physa.2019.01.031.

[34] **Tsallis, C. and Tirnakli, U.** Nonadditive entropy and nonextensive statistical mechanics - Some central concepts and recent applications. Journal of Physics Conference Series 2009. 201(1), arXiv. DOI:10.1088/1742-6596/201/1/012001

[35] **Wilk, G. and Włodarczyk, Z.** Some Non-Obvious Consequences of Non-Extensiveness of Entropy. Entropy. 2023; 25(3):474. https://doi.org/10.3390/e25030474

[36] **Sigalotti, L.D.G.; Ramírez-Rojas, A. and Vargas, C.A.** Tsallis q-Statistics in Seismology. Entropy. 2023; 25(3):408. https://doi.org/10.3390/e25030408

[37] **Tong, S.; Bezerianos, A.; Paul, J.; Zhu, Y. and Thakor, N.** Nonextensive entropy measure of EEG following brain injury from cardiac arrest. Statistical Mechanics and its Applications. Volume 305, Issues 3–4, 2002, Pages 619-628, ISSN 0378-4371, https://doi.org/10.1016/S0378-4371(01)00621-5.

[38] **Zhang, D.; Jia, X.; Ding, H. and et al.** Application of Tsallis entropy to EEG: quantifying the presence of burst suppression after asphyxial cardiac arrest in rats. IEEE Trans Biomed Eng. 2010 Apr;57(4):867-74. doi: 10.1109/TBME.2009.2029082. Epub 2009 Aug 18. PMID: 19695982; PMCID: PMC3050535.

[39] **Thilagaraj, M.; Rajasekaran, P. and Kumar, A.** Tsallis entropy: as a new single feature with the least computation time for classification of epileptic seizures. Cluster Comput 22 (Suppl 6), 15213–15221 (2019). https://doi.org/10.1007/s10586-018-2549-5

[40] **Redelico, F.O.; Traversaro, F.; García, M.D.C. and el al.** Classification of Normal and Pre-Ictal EEG Signals Using Permutation Entropies and a Generalized Linear Model as a Classifier. Entropy. 2017; 19(2):72. https://doi.org/10.3390/e19020072

[41] **Amaral-Silva, H.; Wichert-Ana, L.; Murta, L.O. and et al.** The Superiority of Tsallis Entropy over Traditional Cost Functions for Brain MRI and SPECT Registration. Entropy. 2014; 16(3):1632-1651. https://doi.org/10.3390/e16031632

[42] **Healy, A.; Burgess-Walker, P.; Naemi, R. and Chockalingam, N.** Repeatability of WalkinSense® in shoe pressure measurement system: A preliminary study. The Foot, 2012, 22(1), 35-39. doi:10.1016/j.foot.2011.11.001

[43] **Holleczek, T.; Ruegg, A.; Harms, H. and Tro, G.** Textile pressure sensors for sports applications. 2010 IEEE Sensors. Waikoloa, HI, USA, 1–4 November 2010, pp. 732-737 doi:10.1109/icsens.2010.5690041

[44] **Saito, M.; Nakajima, K.; Takano, C.; Ohta, Y.; Sugimoto, C. and et al.** An in -shoe device to measure plantar pressure during daily human activity. Medical Engineering & Physics. 2011, 33(5), 638-645. doi:10.1016/j.medengphy.2011.01.001

[45] **Salpavaara, T.; Verho, J.; Lekkala, J. and Halttunen, J.** Wireless insole sensor system for plantar force measurements during sport events. Proceedings of IMEKO XIX World Congress on Fundamental and Applied Metrology, Lisbon, Portugal, 2009, 2118–2123.

[46] **Shu, L.; Hua, T.; Wang, Y.; Li, Q.; Feng, D. D. and Tao, X.** In-shoe plantar pressure measurement and analysis system based on fabric pressure sensing array. IEEE Transactions on Information Technology in Biomedicine, 2010, 14(3), 767-775. doi:10.1109/titb.2009.2038904

[47] **Günaydın, B. and İkizoğlu, S.** Multifractal detrended fluctuation analysis of insole pressure sensor data to diagnose vestibular system disorders. Biomed. Eng. Lett. (2023). https://doi.org/10.1007/s13534-023-00285-9

[48] **Tahir, A.M.; Chowdhury, M.E.; Khandakar, A. and et al.** A Systematic Approach to the Design and Characterization of a Smart Insole for Detecting Vertical Ground Reaction Force (vGRF) in Gait Analysis. Sensors, 2020, 20(4), 957. doi:10.3390/s20040957

[49] **Sensor Link:** https://cdn2.hubspot.net/hubfs/3899023/Interlinkelectronics%20November2017/Docs/Datasheet_FSR.pdf (accessed on 1 March, 2023)

[50] **Linardatos, P.; Papastefanopoulos, V. and Kotsiantis, S.** Explainable AI: A Review of Machine Learning Interpretability Methods. Entropy. 2021; 23(1):18. https://doi.org/10.3390/e23010018

[51] **Burden, R.L. and Faires, J.D.** Numerical Analysis. Boston, MA: Cengage Learning, 2019, 144-172.

[52] **Peterson, L.** (2009). K-nearest neighbor. *Scholarpedia, 4*(2), 1883. doi:10.4249/scholarpedia.1883

[53] **Schober, P. and Vetter T.R.** Logistic Regression in Medical Research. Anesth Analg. 2021 Feb 1;132(2):365-366. doi: 10.1213/ANE.0000000000005247. PMID: 33449558; PMCID: PMC7785709.

[54] **Geron, A.** (2019). Chapter 5: Support Vector Machines. In Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc.: Sebastopol, CA, USA, 2019.

[55] **Lin, Y.; Wang, C.; Wu, T.; Jeng, S., and Chen, J.** Support vector machine for EEG signal classification during listening to emotional music. *2008 IEEE 10th Workshop on Multimedia Signal Processing,* Cairns, QLD, Australia, 8–10 Oct 2008, 127-130. doi:10.1109/mmsp.2008.4665061

[56] **Saccà, V.; Campolo, M.; Mirarchi, D.; Gambardella, A.; Veltri, P. and Morabito, F.** On the Classification of EEG Signal by Using an SVM

Based Algorithm; Springer: Cham, Switzerland, 2018, doi:10.1007/978-3-319-56904-8_26.

[57] **Saini, I.; Singh, D. and Khosla, A.** QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. *Journal of Advanced Research, 2013, 4*(4), 331-344. doi:10.1016/ j.jare.2012.05.007

[58] **Yean, C.W.; Khairunizam, W.; Omar, M.I.; Murugappan, M. et al.** Analysis of the distance metrics of KNN classifier for EEG signal in stroke patients. *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA),* Kuching, Malaysia, 15–17 August 2018. doi:10.1109/icassda.2018.8477601

[59] **Erguzel, T.T.; Noyan, C.O. and Eryilmaz, G.** et al. Binomial Logistic Regression and Artificial Neural Network Methods to Classify Opioid-Dependent Subjects and Control Group Using Quantitative EEG Power Measures. Clinical EEG and Neuroscience. 2019;50(5):303-310. doi:10.1177/1550059418824450

[60] **Maria, G.; Juan, S. and Helbert, E.** EEG signal analysis using classification techniques: Logistic regression, artificial neural networks, support vector machines, and convolutional neural networks, Heliyon, 7(6), 2021, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2021.e07258.

APPENDICES

APPENDIX A: Comparative Figures for All Stages of Data Processing and Analysis
APPENDIX B: Matlab Codes for All Stages of Data Processing and Analysis.

APPENDIX A: Comparative Figures for All Stages of Data Processing and Analysis.

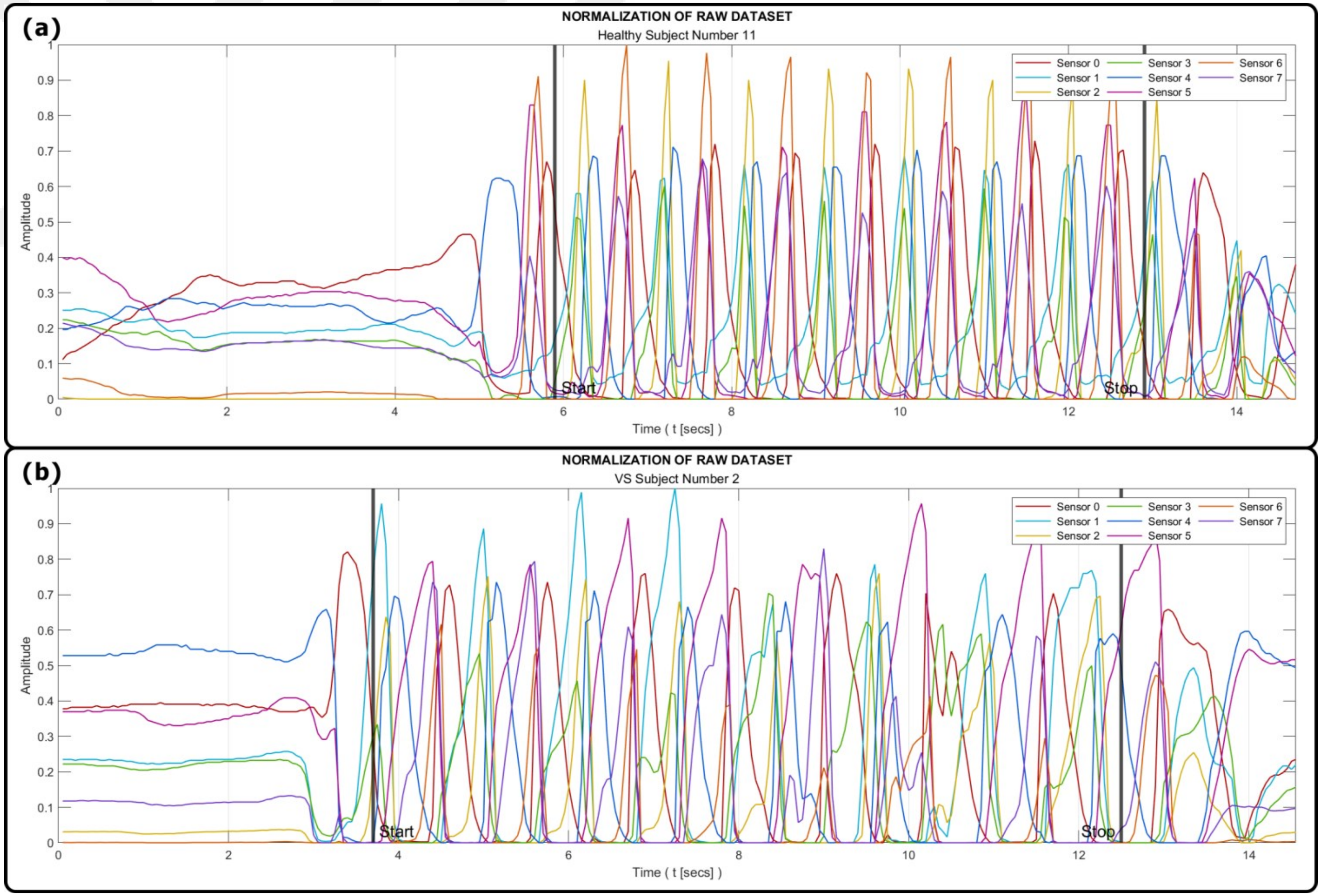


Figure A.1: Normalization and cropping of raw sensor data (Stage 1) for **(a)** healthy and **(b)** subject with vestibular system dysfunction.

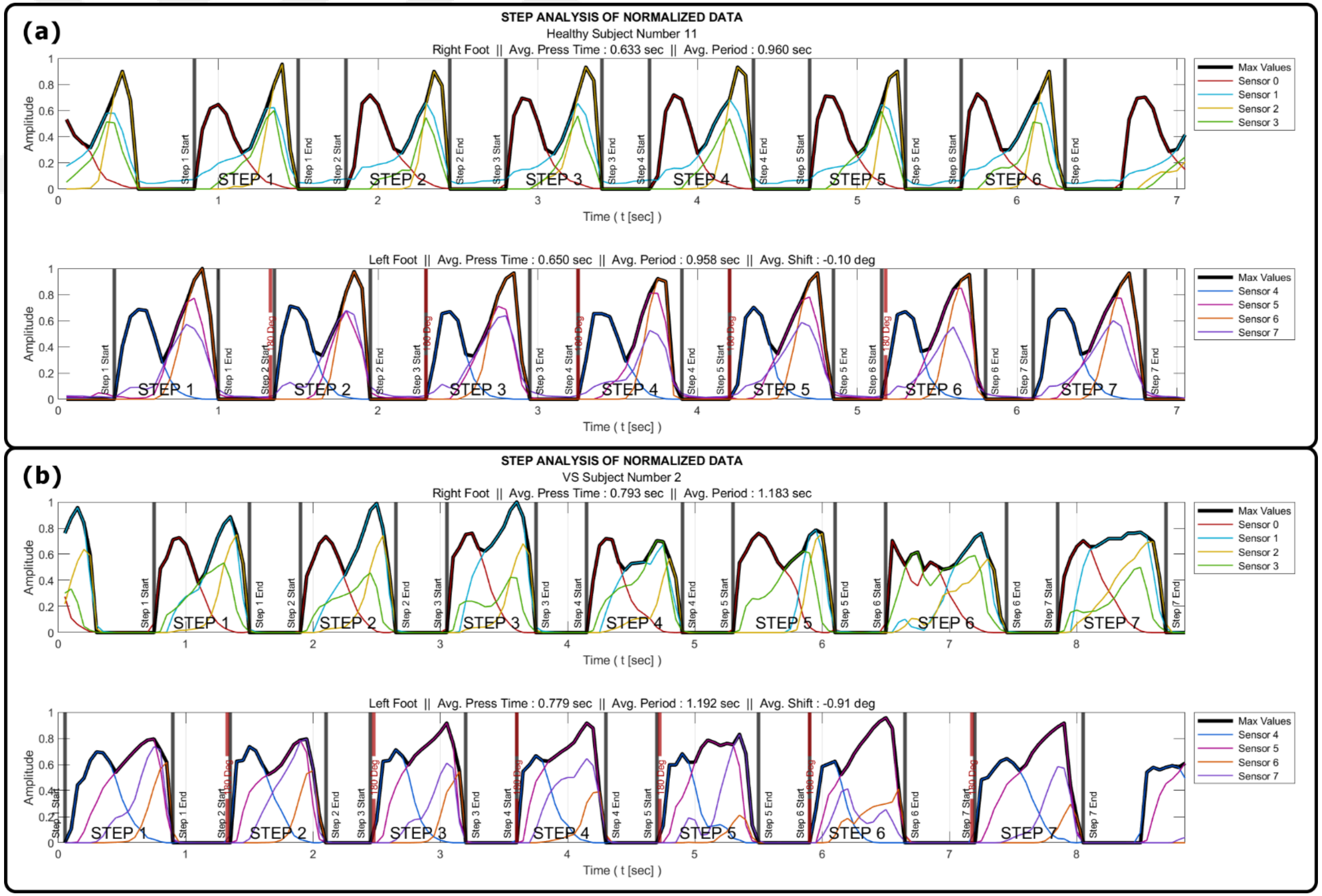


Figure A.2: Step Analysis (Stage 2) of **(a)** healthy and **(b)** subject with vestibular system dysfunction.

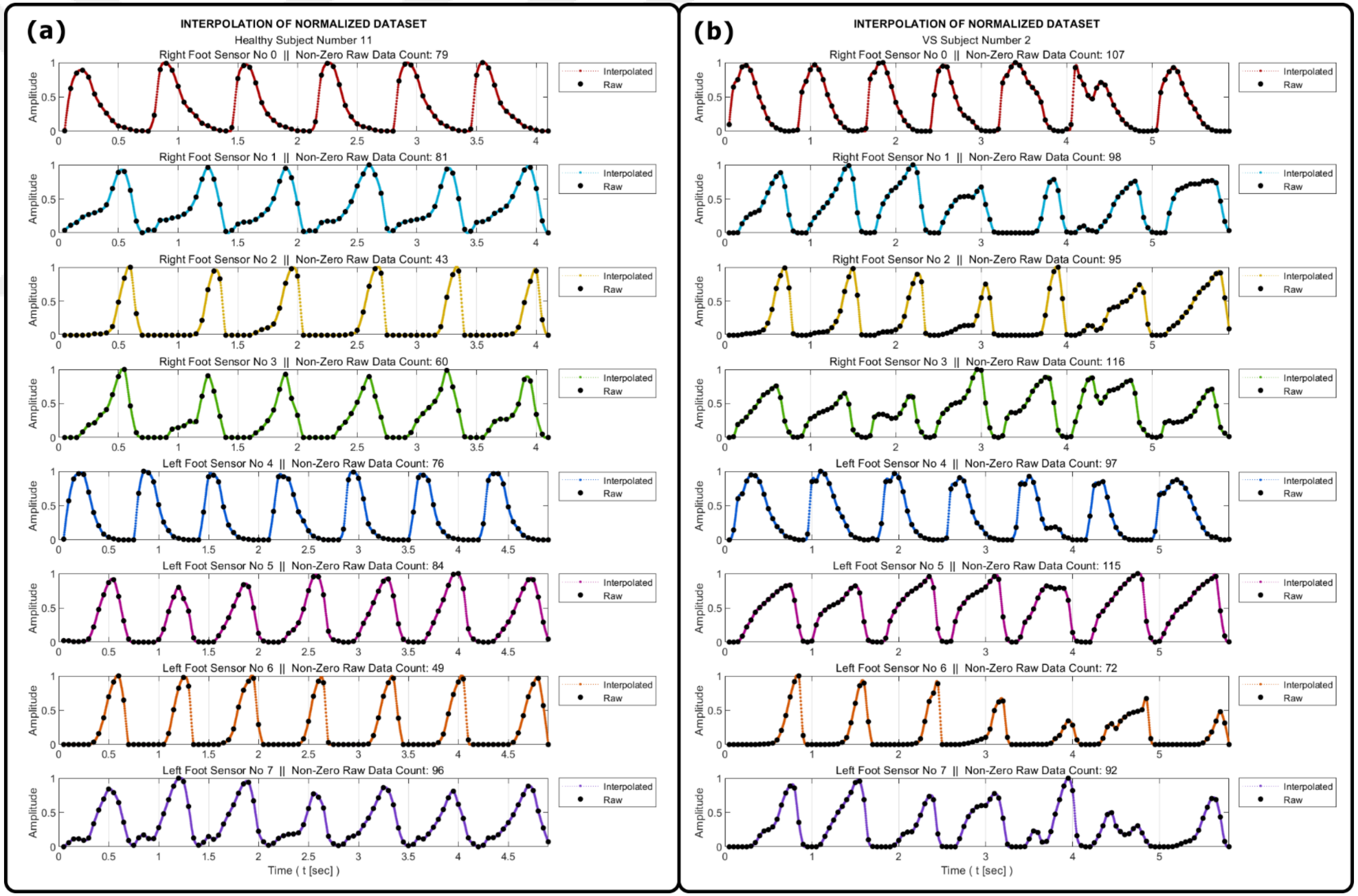


Figure A.3: Interpolation of raw sensor data (Stage 3) of **(a)** healthy and **(b)** subject with vestibular system dysfunction.

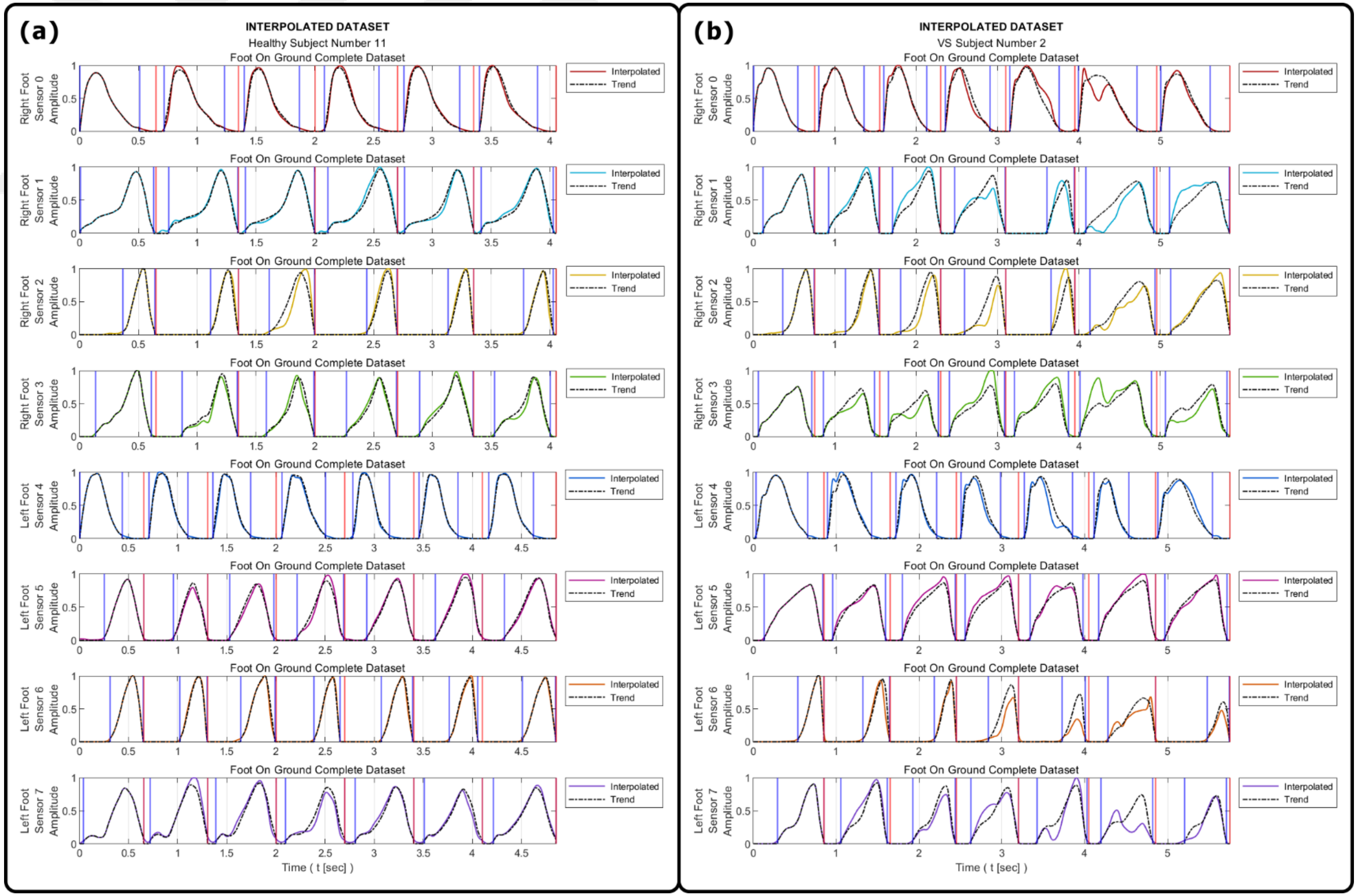


Figure A.4: Trend curve of each step data in the interpolated data (Stage 4) of **(a)** healthy and **(b)** subject with vestibular system dysfunction.

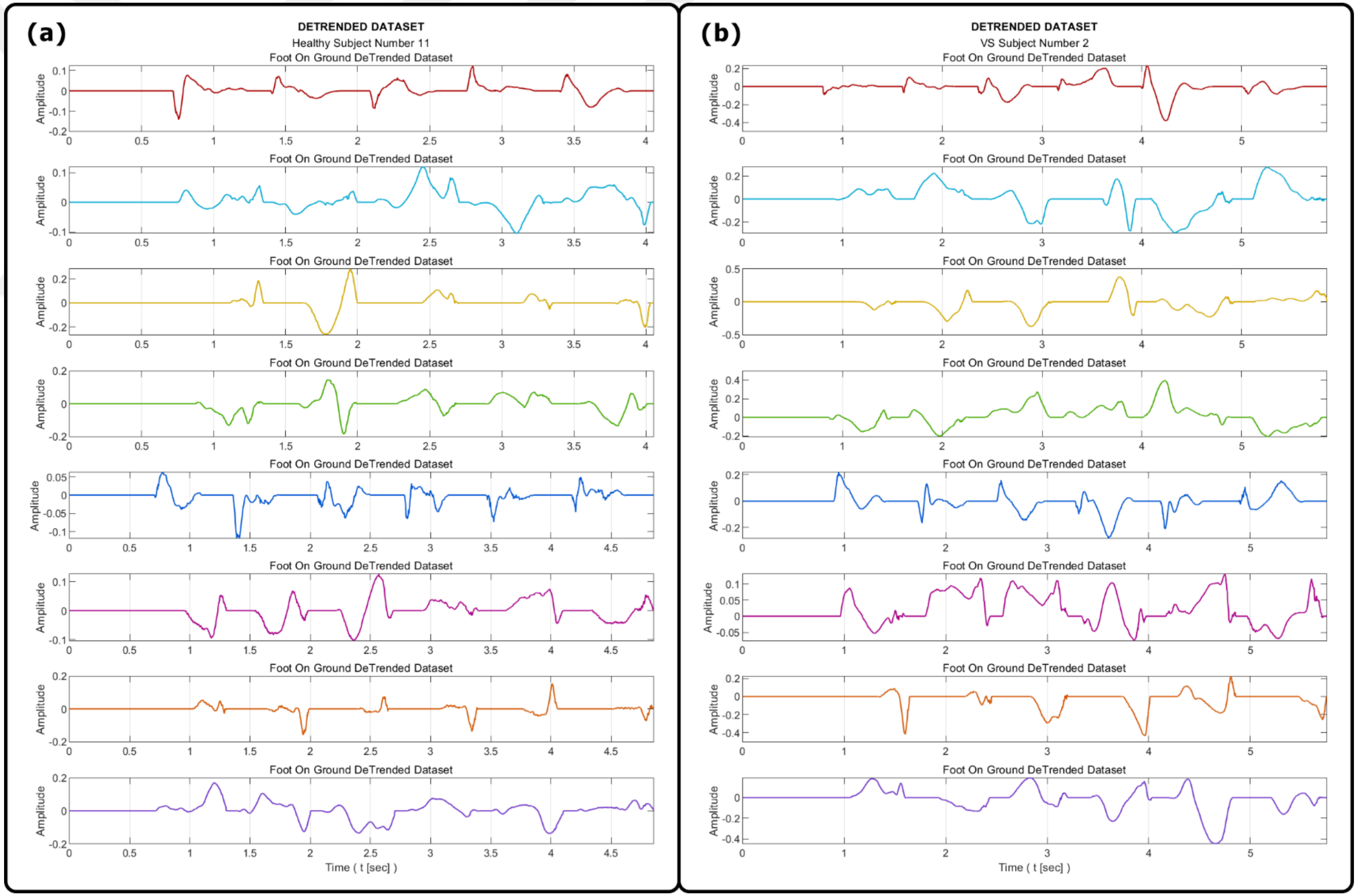


Figure A.5: Obtaining the detrend data (Stage 4) of **(a)** healthy and **(b)** subject with vestibular system dysfunction.

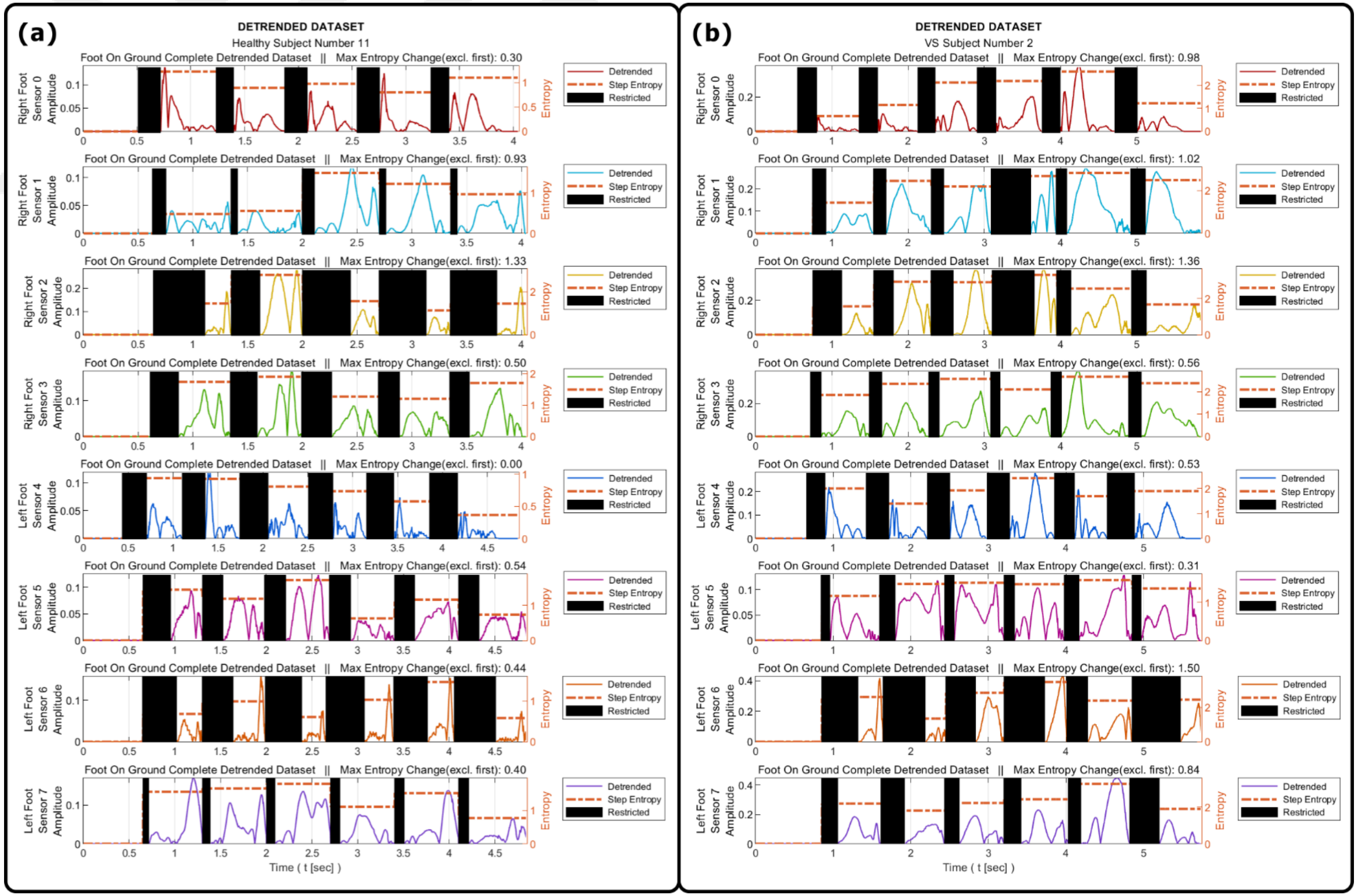


Figure A.6: Absolute value of detrended data and step-wise entropy values (Stage 5) of **(a)** healthy and **(b)** subject with vestibular system dysfunction.

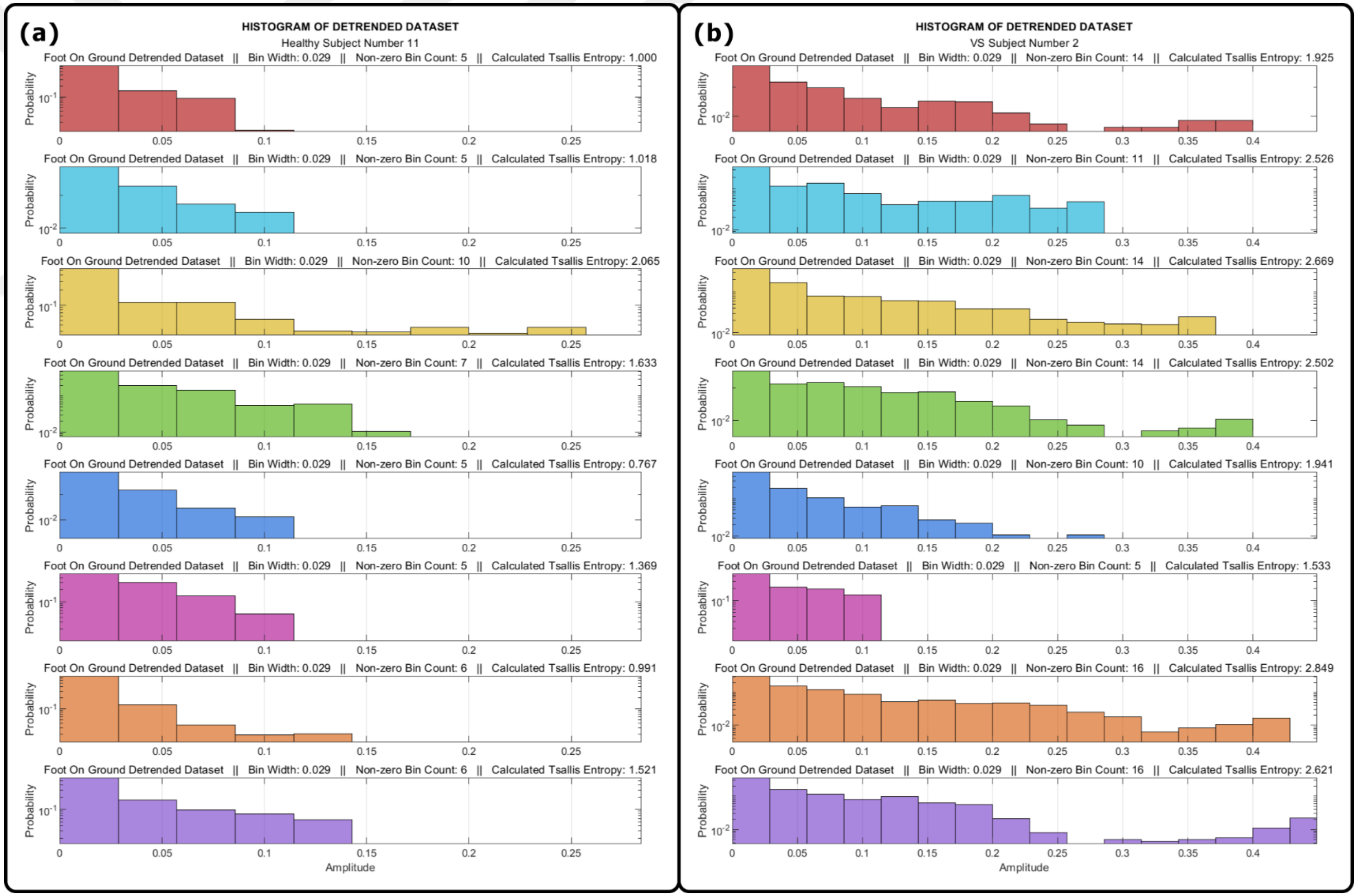


Figure A.7: Histogram graph and total entropy values (Stage 5) of **(a)** healthy and **(b)** subject with vestibular system dysfunction.

APPENDIX B: Matlab Codes for All Stages of Data Processing and Analysis.

Stage 1 – Data Normalization and Cropping

The primary objective of this document is to standardize the accessible data by removing unsuitable sections for gait analysis. It involves visualizing these procedures and generating normalized and cropped datasets as output.

▶ A) Importing Raw Datasets

▶ B) Important Variables

▶ C) Trimming & Plotting of Raw Data

▼ D) Processing of Data

Utilizes eight time-series sensor data collected from the designated participant. Firstly, the dataset is normalized within the range of 0 to 1. Subsequently, the segments relevant for gait analysis are identified and annotated within the datasets. Notifications are generated for any sensor data that could not be obtained. These procedures result in the creation of .png files, which are stored in individual folders corresponding to each participant. The normalized and cropped datasets are then consolidated into a single .mat file, which is saved in the designated folder.

- **Dataset_:** Raw datasets of all participants as type of struct.
- **Frames_:** All Frames as type of cell array, specifying sections that are suitable for data analysis.
- **HEVS:** Indication of Healthy or VS participants. 1: Healthy 2: VS
- **No:** Participant Number 1 to 30.
- **Hide:** Hide participants' names in plots.
- **[NDS]:** Normalized & cropped dataset of related participant.

```
function [NDS] = ProcessData(Dataset_, Frames_, HEVS, No)
...
frames = cell2mat(Frames_(HEVS, No));
...

% Normalization to range of 0 - 1
min_val = min(min(Data));
Data = Data - min_val;
max_val = max(max(Data));
Data = Data / max_val;
...

% Cropping the dataset from the frame boundaries
NDS = Data(:, Frames(1):Frames (2));
...

end
```

Stage 2 – Step Analysis

The purpose of this file is to perform the necessary pre-processing steps on the sensor data prior to interpolation. These steps include separating the sensor data into left and right feet, identifying the start and end points of each step, and capturing various attributes such as step duration and foot pressure duration.

▶ A) Importing Normalized Datasets

▶ B) Important Variables

▶ C) Step Analysis

▼ D) Processing of Data

In this section, the sensor data is initially segregated based on the feet. Subsequently, the fixed data sensors are identified. Next, the individual steps taken by each foot are determined. The push time, repetition time, and position of these steps within the normalized dataset are computed. Finally, all of these procedures are visually depicted and saved in a designated folder for inspection.

- **Dataset_:** Normalized datasets of all participants as type of struct.
- **Threshold_:** Threshold value for detecting abnormal steps.
- **HEVS:** Indication of Healthy or VS participants. 1: Healthy 2: VS
- **No:** Participant Number 1 to 30.
- **Plot_:** Plot and export result. (Significantly affects processing time.)
- **[info]:** contains the positions of the starting and ending points of right and left steps in the normalized dataset. It also includes calculated values such as step times, periods…

```
function [info] = ProcessData(Dataset_, Threshold_, HEVS, No, plot_)
th_ns_r = Threshold_(1); % Non-step treshold value right
th_ns_l = Threshold_(2); % Non-step treshold value left
...
% ~~~~~~~~~~~~ RIGHT FOOT ~~~~~~~~~~~~~~~~
% Combining four sensor data of a foot with the maximum operator.
maxin = max(Data(1:4,:));
...
% A series of loop operations to detect steps.
for i = 2:length(maxin)
      ...
end
...
% Contact Time and Step Period Calculation
Contact_Time_Right = Step_Right(2,:) - Step_Right(1,:);
for i = 2:length(Step_Right(1,Step_Right(1,:)>0);)
    Period_Right(i-1) = Step_Right(1,i) - Step_Right(1,i-1);
end
...
end
```

Stage 3 – Interpolation

The objective of this file is to merge the step data acquired in the preceding stage and execute the interpolation of said data. Since these procedures are conducted for every participant, it may take approximately 5-10 minutes. Consequently, we will carry out the creation of the histogram in the subsequent phase.

▶ A) Importing Normalized Datasets

▶ B) Important Variables

▶ C) Interpolation of Sensor Data

▼ D) Processing of Data

In this section, the sensor data is initially segregated based on the feet. Subsequently, the fixed data sensors are identified. Next, the individual steps taken by each foot are determined. The push time, repetition time, and position of these steps within the normalized dataset are computed. Finally, all of these procedures are visually depicted and saved in a designated folder for inspection.

- **Dataset_:** Normalized & Clipped datasets with step information of all participants as type of struct.
- **HEVS:** Indication of Healthy or VS participants. 1: Healthy 2: VS
- **No:** Participant Number 1 to 30.
- **Method:** Interpolation method.
- **Coeff:** Resolution enhancement coefficient.
- **Plot_:** Plot and export result. (Significantly affects processing time.)
- **[Processed_Dataset]:** Output containing the interpolated and non-interpolated right and left foot sensors added one after the other when the foot touches the ground.

```
function [Right_Step, Left_Step] = ProcessData(Dataset_, HEVS, No,
Method, Coeff, plot_)
rsl = 0.05; % Raw Data Resolution [sn/data]
cof = 1/Coeff; % Interpolated Data Cooefficent. It means new resolution
is rsl*cof [sn/data]
...
t = 1:total_length;
t = t*rsl;
tnew = 1:cof:total_length;
tnew = tnew*rsl;
inter = interp1(t, AllSteps, tnew, Method);
...
end
```

Stage 4 – Detrending of Interpolated Dataset

The primary objective of the present document is to eliminate the inherent trend observed within the interpolated data. Throughout the execution of this procedure, the ultimate goal is to visually represent the computed polynomial for every iterative step undertaken.

▶ **A) Importing Interpolated Datasets**

▶ **B) Important Variables**

▶ **C) Detrending of Interpolated Data**

▼ **D) Processing of Data**

In this section, firstly, all the information and dataset obtained in the previous stages are imported. In the initial step, the interpolated data is reconstructed using a simplified approach. In the second step, a threshold level for this dataset is determined, and a quadratic detrend polynomial is calculated for each fluctuation step within the dataset. These detrend polynomials are then plotted on the graph. Subsequently, the detrended data is visualized in its new form, and the output of all the work is saved in the designated folder.

- **Dataset_:** All the information & dataset obtained in the last stages.
- **HEVS:** Indication of Healthy or VS participants. (1:Healthy, 2:VS)
- **No:** Participant Number 1 to 30.
- **Affinity:** Trend variation coefficient - affinity coefficient.
- **Thresh_:** Degree of trend polynomial equation.
- **Plot_:** Plot and export result. (Significantly affects processing time.)
- **[StepR_Data_S2]:** Detrended Dataset of Right Foot.
- **[StepL_Data_S2]:** Detrended Dataset of Left Foot.

```
function [StepR_Data_S2, StepL_Data_S2] = ProcessData(Dataset_, HEVS,
No, Affinity, Thresh_, Plot_)
...
Trend = imresize(Last_Trend_Curve, [1 length(fluctuation)], 'nearest');
old_error = mean(abs(fluctuation - Trend)); step = 0.001; found = 0;
yon = 0; Distance = 0;
while ~found
   ...
   predicted_trend = Trend * (1 - Distance) + C * Distance;
   error = mean(abs(fluctuation - predicted_trend));
   if error > old_error; yon = ~yon; step = step/2; end
   old_error=error;
   if abs(Distance) >= aff_coef || step < 0.00001; found = 1;  end
end
Last_Trend_Curve = predicted_trend; Trend = predicted_trend;
Detrend = fluctuation - Trend;
...
end
```

Stage 5 – Entropy Calculations

The primary objective of this document is to generate a graphical representation, specifically a histogram, of the data that has undergone both detrending and interpolation processes. Subsequently, the calculation of Tsallis entropy will be carried out on the aforementioned data.

▶ **A) Importing Dataset**

▶ **B) Important Variables**

▶ **C) Histogram and Entropy Study**

▼ **D) Processing of Data**

In this section, initially, all the information and dataset obtained in the preceding stages are imported. In the first stage, the detrended interpolated data is regraphed using a straightforward approach. In the second stage, histograms of these data are analyzed and plotted. Finally, the entropy calculation is executed, and the visual representation of the analysis is saved to the designated folder.

- **Dataset_:** All the information & dataset obtained in the last stages.
- **HEVS:** Indication of Healthy or VS participants. 1: Healthy 2: VS
- **No:** Participant Number 1 to 30.
- **Bin_Width:** Bin width of histogram.
- **Method:** Entropy calculation method.
- **Param:** Parameter of entropy calculation method if there is any.
- **Plot_:** Plot and export result. (Significantly affects processing time.)
- **[Probability]:** Probability values in the Histogram.
- **[Entropy]:** Calculated entropy values.
- **[EES]:** Entropy values for each steps.

```
function [Probability, Entropy, EES] = ProcessData(Dataset_, HEVS, No,
Bin_Width, Method, Param, Plot_)
...
subplot(8,2,2*i);
Hs = histogram(Detrended_Data, rsl_his, Normalization="probability",
FaceColor=colors_(i), EdgeColor='k', BinWidth = 1/rsl_his);
Probability(i,1:length(Hs.Values)) = Hs.Values; % Probabilities
...
Entropy(i,1) = tsallisEntropy(Probability,Param);
...
end

function entropy = tsallisEntropy(probabilities, parameter)
...
q = parameter; % Tsallis parameter
entropy = (1 - sum(Prob_nonzero.^q)) / (q - 1);
...
end
```

Stage 6 – Feature Extraction

In order to facilitate the process of feature extraction for machine learning, employing the acquired data is imperative. This entails identifying and isolating relevant attributes or characteristics from the data set that can serve as inputs for subsequent machine learning algorithms.

▶ A) Importing Dataset

▶ B) Feature Extraction

▼ C) Processing of Data

In this section, firstly, all the information & dataset obtained in the last stages is imported. For a better understanding of the obtained data, operations such as standard deviation are applied. It is then tabulated and clustered as inputs/outputs.

- **Dataset_:** All the information & dataset obtained in the last stages.
- **[Inputs]:** Inputs for machine learning.
- **[Outputs]:** Outputs for machine learning.

```
function [Inputs, Outputs] = ProcessData(Dataset_)
...
 for p = 1:2 % Min-Max 1:2
   for q = 1:30  % Min-Max 1:30
    if p == 1
        Ent = Dataset_.Healthy(q).Entropy';
        SSER = cell2mat(Dataset_.Healthy(q).EES(1));
        SSEL = cell2mat(Dataset_.Healthy(q).EES(2));
    else
        Ent = Dataset_.VS(q).Entropy';
        SSER = cell2mat(Dataset_.VS(q).EES(1));
        SSEL = cell2mat(Dataset_.VS(q).EES(2));
    end
    for r = 1 : 4; StSER(30*(p-1) + q,r   ) = Evaluate(SSER(r  ,:)); end
    for r = 5 : 8; StSEL(30*(p-1) + q,r-4) = Evaluate(SSEL(r-4,:)); end
    Total_Entropy(30*(p-1) + q ,:) = Ent;
   end
 end
...
End

function [result] = Evaluate(SSE_)
...
SSE_(1) = []; % Removing entropy value of first step
result = std([SSE_ -SSE_]); % Calculation of deviation from zero
end
```

CURRICULUM VITAE

Name Surname : Harun Yaşar KÖSE

EDUCATION:

- **B.Sc.** : 2015, Kocaeli University, Engineering Faculty, Department of Mechatronics Engineering

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Köse, H.Y.**; İkizoğlu, S. Development of An Entropy-Based Method for Determining the Balance Problem in Individuals, 8^{th} International Battalgazi Scientific Studies Congress, Sep 3, 2023 Malatya, p14-24. ISBN: 978-625-367-296-6 **(Congress Presentation & Publication)**
- **Köse, H.Y.**; İkizoğlu, S. Nonadditive Entropy Application to Detrended Force Sensor Data to Indicate Balance Disorder of Patients with Vestibular System Dysfunction. Entropy 2023, 25, 1385. https://doi.org/10.3390/e25101385 **(Article)**