sEMG-based Predictive Framework for Injury Risks in Sports: Novel Deep Learning Approaches Tested on EPL Athletes

Ergeneci, Mert

Awarding institution:
King's College London

sEMG-based Predictive Framework for Injury Risks in Sports: Novel Deep Learning Approaches Tested on EPL Athletes

Mert Ergeneci

Supervisor: Dr. Panagiotis Kosmas, 1st Supervisor
Prof. Osvaldo Simeone, 2nd Supervisor

The Department of Engineering
King's College London

This dissertation is submitted for the degree of
Doctor of Philosophy

July 2024

I would like to dedicate this thesis to my loving wife Duygu …

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 100,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Mert Ergeneci
July 2024

Acknowledgements

First and foremost, I owe a deep debt of gratitude to my primary supervisor, Dr. Panos Kosmas. His unwavering support, inspiration, and guidance from the inception to the completion of this research have been invaluable. His keen understanding of my mentality and approach greatly enhanced the feedback and constructive criticism I received, all of which significantly contributed to the quality of this thesis. A heartfelt thank you to him for making this journey both possible and rewarding.

My sincere appreciation extends to my second supervisor, Prof. Osvaldo Simeone. His expertise, particularly in the domains of deep learning and meta learning, enriched my research approach. His invaluable feedback and insights have been instrumental in refining my work.

I wish to extend my profound thanks to Manchester United's dedicated staff, particularly David Binningsley, whose relentless assistance in data collection and supervision regarding sports science aspects has been pivotal. His dedication, passion, and expertise have been indispensable to this research. My gratitude also goes to Michael Clegg for his invaluable strength and conditioning advice, as well as his contributions to data collection. I am grateful to Robin Saddler for the collaborative efforts, and a special mention to Richard Hawkins for his unwavering support.

At Leeds United, Daryl Carter stands out for his exceptional consultancy, assistance with data collection, and innovative supervision on sports science aspects. His commitment and enthusiasm have greatly elevated the quality of my work. I would also like to express my gratitude to Rob Price for his contributions and unwavering support.

I am thankful to the team at Neurocess for providing the sensors and systems that were integral to this research. Their commitment to innovation and research has been truly commendable.

In conclusion, to all those who contributed, both directly and indirectly, to this thesis — I am forever indebted to you. A special gratitude to my family for their unwavering love and support. Your collective belief and encouragement have been the foundation of this work.

Abstract

Surface electromyography (sEMG) measures electrical signals in muscles during contraction. Its established applications include medical diagnostics, prosthetics, and rehabilitation. However, in sports science, particularly professional football (soccer), sEMG's potential remains largely unexplored. This research investigates the use of sEMG in football, which is significantly affected by both financial implications and competitive performance due to injuries. The study focuses on commonly observed football injuries: hamstring, adductor, and soleus injuries (HASI). The analysis utilises sEMG data from 45 elite athletes affiliated with two top English Premier League (EPL) clubs over 2.5 seasons. Data collection, predominantly on the day following matches (MD+1), aligns with professional sports practices, enhancing the dataset's relevance for injury analysis.

The core novelty of this thesis is a comprehensive framework that transforms raw sEMG data into actionable injury predictions. This systematic approach meticulously converts sEMG recordings into probabilities of HASI occurrence, with each stage contributing distinctly to sports science and sEMG analytics.

A significant innovation is the development of a deep learning model for accurate drill-time identification in continuous sEMG recordings. This model effectively discerns genuine drill-related muscle contractions, addressing a common challenge in continuous data segmentation within sports. Athletes in sports settings undertake various activities such as walking, jumping, or stretching during exercise interludes. This adds complexity in differentiating exercise-related muscle activations from incidental movements. The drill-time labelling model incorporates custom loss regularisation specific to football training and a kernel and depth selector attention module. This module enhances feature learning versatility and model performance, accurately identifying drill times amid diverse athletic activities.

Furthermore, the study introduces a novel approach for denoising Motion-Induced Artefacts (MIA) in sEMG data, which is common in dynamic sports environments. Athletes engaging in high-motion exercises result in unavoidable MIA inclusion. The proposed model first detects contaminated zones in the data sequence and then employs a deep learning-based denoising process to clean the data while preserving the original sEMG signals' integrity. This dual approach outperforms existing state-of-the-art models and is particularly effective in sports settings.

In motion classification, the research surpasses existing models on public hand gesture datasets and adeptly categorises football-specific lower extremity drills. The deep learning strategy divides input sEMG data into low and high-frequency bands, processed in parallel. Channel and temporal attention mechanisms focus on significant features within each band, enhancing performance and versatility. The model is effective for both hand gesture recognition and lower extremity football drills without

parameter tuning. Additionally, a novel few-shot metric learning approach enhances classification accuracy, especially in sports settings with limited data. This approach outperforms state-of-the-art few-shot learning applications, demonstrating effectiveness in sparse dataset scenarios.

Another key innovation is the injury prediction model, hypothesising that muscle recruitment pattern deviations indicate impending injuries. The study uses a meta-learning-based deep metric learning technique to detect these deviations. This technique learns a representation space where healthy athlete states are near their baseline centroid, and risky states are more distant, forming a cluster between post-injury and safe data points. The model efficiently detects anomalies and accurately predicts HASI injuries, rapidly adapting to individual athletes' risk traits. This adaptability is vital, considering that each athlete's risk factors related to physiological states are unique and subjective. By learning these individual characteristics effectively, the model offers a nuanced and athlete-specific understanding of injury risks, enhancing the precision and relevance of HASI predictions.

In conclusion, this research underscores the substantial potential of combining sEMG with deep learning in sports science, particularly for injury prevention in professional athletes. It integrates comprehensive datasets, novel methodologies, and practical applications, positioning itself as a pivotal reference for future developments in sports injury prediction and prevention.

Table of contents

List of figures **xi**

List of tables **xviii**

1 Introduction **1**
1.1 Overview and Motivation . . . 1
1.2 Introduction To The Problem of Interest . . . 2
1.2.1 The Cost of Injury . . . 2
1.2.2 Prevalence of Muscle Injuries in Football . . . 2
1.3 Practical Application of Injury Prevention Framework and Data Collection . . . 3
1.4 The Proposed Injury Prevention Framework . . . 4
1.5 Challenges of sEMG Utilisation in Sports Settings . . . 5
1.6 Contribution and Relevant Publications . . . 8
1.7 Thesis Structure . . . 10

2 Background **11**
2.1 Muscle Injuries: Classification and Mechanisms . . . 11
2.1.1 Hamstring: Anatomy and Injury Mechanisms . . . 13
2.1.2 Adductor: Anatomy and Injury Mechanisms . . . 14
2.1.3 Soleus: Anatomy and Injury Mechanisms . . . 14
2.2 Surface Electromyography (sEMG): An Overview . . . 15
2.2.1 Mechanisms of sEMG Signal Generation . . . 15
2.2.2 Characteristics and Factors Influencing the sEMG Signal . . . 16
2.2.3 Fundamental Approaches to sEMG Signal Analysis . . . 17
2.3 Foundations of Artificial Neural Networks: An Overview . . . 21
2.3.1 Understanding Deep Learning Processes: Training, Inference, and Testing . . . 22
2.3.2 Loss Functions and Optimisers in Deep Learning . . . 23
2.3.3 Neural Networks Layers . . . 26
2.3.4 Introduction to Performance Metrics . . . 32

3 Datasets Utilised: Acquisition, Description, and Experimental Context **34**
3.1 Ninapro Dataset (D_n) . . . 34
3.2 Football Dataset (D_f) . . . 35

3.2.1 Optimal sEMG Sensor Placement . . . 35
3.2.2 Description of Selected Drills . . . 36
3.3 Comparative Analysis of D_n and D_f: Suitability for Thesis Scope . . . 39

4 Drill-Time Labelling **47**
4.1 Introduction . . . 47
4.2 State-of-the-art . . . 48
4.3 Contributions . . . 48
4.4 Proposed Methodology . . . 49
4.4.1 Preprocessing . . . 49
4.4.2 Network Structure . . . 49
4.4.3 Spike Loss Regularisation . . . 54
4.5 Experiments . . . 55
4.6 Results . . . 56
4.7 Conclusion . . . 57

5 Motion-Induced Artefact (MIA) Detection and Denoising **61**
5.1 Introduction . . . 61
5.2 State-of-the-art . . . 62
5.3 Contributions . . . 62
5.4 Problem Definition . . . 63
5.5 Proposed Methodology for MIA Detection . . . 63
5.6 Onset Detection with MIA in sEMG Streams . . . 65
5.6.1 State-of-the art onset detection methods . . . 65
5.6.2 Experiments and Results . . . 67
5.7 Identifying MIA-Contaminated Zones in sEMG Regions . . . 70
5.7.1 Synthetic MIA Model . . . 71
5.7.2 Contamination of D_n with Synthetic MIA . . . 72
5.7.3 Experiments and Results . . . 73
5.8 Proposed U-Net Encoder Decoder (UED) Model for Denoising MIA from sEMG Signals 75
5.9 Validation Experiments for the Denoising Framework . . . 77
5.10 Results . . . 78
5.11 Conclusion . . . 80

6 Motion Classification **81**
6.1 Deep Learning Model for Motion Classification . . . 81
6.1.1 Introduction . . . 81
6.1.2 State-of-the-art . . . 82
6.1.3 Contributions . . . 82
6.1.4 Proposed Model . . . 83
6.1.5 Experiments & Results . . . 85

6.2 Adapting Deep Learning for Constrained Datasets . . . 90
6.2.1 Introduction . . . 90
6.2.2 State-of-the-art . . . 90
6.2.3 Contributions . . . 91
6.2.4 Dataset & Preprocessing . . . 91
6.2.5 Few-Shot Metric Learning . . . 92
6.2.6 State-of-the-art Metric Learning (G_ϕ) . . . 93
6.2.7 Proposed Methods . . . 95
6.2.8 Experiments . . . 98
6.2.9 Results . . . 98
6.2.10 Discussions . . . 99
6.3 Conclusion . . . 100

7 Hamstring, Adductor, Soleus Injury (HASI) Prediction 101
7.1 Introduction . . . 101
7.2 Review of sEMG Responses to Muscular Injuries . . . 102
7.3 Selection of sEMG-based Features As Risk Factors . . . 103
7.3.1 Cumulative Recruitment (CR) . . . 103
7.3.2 Compensation . . . 104
7.3.3 Bilateral Similarity (BS) . . . 105
7.3.4 Target Muscle Imbalance (Im) . . . 106
7.3.5 Waveform Peak Detection . . . 107
7.3.6 Explosiveness (V/s) . . . 109
7.3.7 Repeatability of sEMG Measurements Across Various Screenings . . . 109
7.3.8 Validation of Compensation & Bilateral Similarity Metrics . . . 110
7.4 HASI Prediction via Anomaly Detection with Independent Univariate Gaussian Distributions (UGAD) . . . 112
7.5 HASI Prediction via Deep Metric Learning Based Anomaly Detection (DML) . . . 114
7.5.1 Network Structure . . . 114
7.5.2 Training the Model: Loss Function Analysis . . . 114
7.5.3 Injury Prediction through Model Inference . . . 117
7.6 Experiments . . . 118
7.7 Results . . . 121
7.8 Discussion . . . 122
7.9 Conclusion . . . 123

8 Conclusion and Future Work 124
8.1 Summary of Contributions . . . 124
8.2 Limitations . . . 125
8.2.1 Absence of Inertial Measurement Unit (IMU) Sensors . . . 125
8.2.2 Scope and Validation Limitations in Exercise Classification . . . 125

8.2.3 Simplification of Denoising by Synthetic MIA Profile 126
8.2.4 Limitations in Injury Risk Labeling 126
8.2.5 Impact of Player Transfers on Dataset Continuity 126
8.3 Future Methodological Work 127
8.3.1 Utilising GAN for Noise Contamination 127
8.3.2 Motion Classification Test on Different Datasets 127
8.3.3 Advancing Injury Prevention with Synergistic Methodologies 127
8.3.4 Enhancing Centroid Calculation: Adopting Deep Learning Over Arithmetic Mean 127
8.3.5 Incorporating Multiple Drills Simultaneously into HASI Prediction Model . 128
8.3.6 Incorporating Temporal Variations into HASI Prediction Model 128
8.4 Future Application Work 129
8.4.1 Broadening the Scope of HASI 129
8.4.2 Tailoring Predictions for Female Athletes 129
8.4.3 Return-to-Play Decisions Following Rehab 129

References **130**

Appendix A Details of sEMG Data Acquisition Systems **143**
A.1 Neurocess sEMG System Specifications 143
A.2 Comparison with Other sEMG Systems 144

List of figures

1.1 Schematic of the sEMG-based injury prevention framework. The flowchart details the process sequence beginning with sEMG recording, followed by Drill-time Labelling, MIA Denoising, and Motion Classification. The output from Motion Classification is then input into the HASI Prediction model, which calculates the HASI likelihood. The dashed outline denotes the entire framework, which processes the sEMG stream to yield a HASI probability output. Models within the framework are indicated by double-bordered squares. 4

2.1 Illustration of minor and moderate muscle tears on a muscle diagram, sourced from [119]. 12
2.2 Illustration of hamstring muscle groups and associated injury types, sourced from [119]. 13

3.1 Illustration of an athlete performing a hamstring claw exercise. sEMG sensors, attached bilaterally on four muscle groups using adhesive medical tapes, are shown. In this figure, only the sensors on the left soleus and hamstring muscles are visible. Sensors on the biceps femoris long head and semitendinosus are marked with a dashed black line, while the sensor on the soleus is indicated by a straight black line. 37
3.2 The illustrated plot is a segment of an sEMG recording from the dataset D_f, where the sEMG stream is represented in black, displaying both contraction phases and subsequent rest periods. The x-axis represents time, while the y-axis indicates the amplitude. Overlaying the sEMG data is a label line in red, indicating the true drill times as labeled by physiotherapists using the Neurocess software. Notably, in the offset region of the plot, there is an appearance of MIA, which poses a potential challenge for onset detection. This is due to its higher peak-to-peak voltage compared to the muscle contractions, underscoring the complexities involved in accurately interpreting sEMG signals in the presence of such artefacts. 42

3.3 Plots (a) and (b) display sEMG recordings from the D_f dataset, featuring different athletes and muscle groups, and illustrating the contamination of MIA on muscle contraction signals. In both plots, the x-axis represents time, and the y-axis indicates the amplitude in millivolts. Areas of MIA contamination are highlighted within red dashed rectangles. Plot (a) features two sequential MIAs during an isometric contraction of the soleus muscle. In contrast, plot (b) showcases an isotonic contraction of the biceps femoris long head muscle, also affected by MIA. These visualisations, drawn from D_f, aid in understanding the impact of MIA on sEMG signals under varying muscular activities. 43

3.4 Plot of recorded sEMG data from the knife grasping motion in the D_n dataset. The x-axis indicates time, where each index corresponds to 0.5 ms, consistent with a 2 kHz sampling rate. The y-axis measures signal amplitude in millivolts. This plot showcases 12 channels (labelled as ch), each associated with a distinct muscle. Channels 1-8 (ch-1 to ch-8) are placed around the forearm at the radio-humeral joint, capturing a range of muscle signals. Channels 9 and 10 (ch-9 and ch-10) are dedicated to the primary activity of the flexor and extensor digitorum superficialis muscles, respectively. Channels 11 and 12 (ch-11 and ch-12) target the primary activity areas of the biceps brachii and triceps brachii muscles. 44

3.5 Plot of recorded sEMG data from the hamstring claw exercise in the D_f dataset. The x-axis indicates time, with each unit representing 1 ms, corresponding to a 1 kHz sampling rate. The y-axis displays signal amplitude in millivolts. Data are collected using an 8-channel configuration (labelled as ch in the plot), tailored for lower limb muscles: Channels 1-2 (ch-1, ch-2) for adductor longus (left and right), Channels 3-4 (ch-3, ch-4) for biceps femoris long head (left and right), Channels 5-6 (ch-5, ch-6) for soleus (left and right), and Channels 7-8 (ch-7, ch-8) for semitendinosus (left and right). 45

3.6 Plot of recorded sEMG data from the SL elevated glute bridge exercise in the D_f dataset. The x-axis indicates time, with each unit representing 1ms, corresponding to a 1 kHz sampling rate. The y-axis displays signal amplitude in millivolts. Data are collected using an 8-channel configuration (labelled as ch in the plot), tailored for lower limb muscles: Channels 1-2 (ch-1, ch-2) for adductor longus (left and right), Channels 3-4 (ch-3, ch-4) for biceps femoris long head (left and right), Channels 5-6 (ch-5, ch-6) for soleus (left and right), and Channels 7-8 (ch-7, ch-8) for semitendinosus (left and right). 46

4.1 An illustration of the network structure. The system takes an input of 16 subwindows, each with a shape of $8,512$, indicating the channel dimension and the sequence length, respectively. Each of these subwindows undergoes parallel processing in CNN blocks. Subsequently, the generated feature vectors from the CNN blocks are directed to a four-layer deep bidirectional LSTM RNN. Each LSTM instance's output is directed into a fully connected layer with a sigmoid activation function, returning a singular output. The completion of this process results in a predictive sequence with a shape of $(16,1)$, denoting a distinct prediction for every subwindow. 50

4.2 Schematic diagrams of KD block is provided. KD block consists of an initial 1D convolution, parallel D blocks with varying kernel sizes, concatenation with a residual tensor, attention block, and summation with another residual, transforming an input tensor of shape (B,C,L) into an output tensor of shape $(B,512,L)$. 52

4.3 Schematic representation of the D block provided. The D block features cascading 1D Separable Convolution blocks (SConvs) commencing from an input size of $(B,32,L)$. After all the SConv blocks are processed, their outputs are merged, tripling the channel dimension to produce the concluding shape of $(B,96,L)$. Combined, these structures refine feature learning by leveraging multiple depths and kernel dimensions, enhancing the overall efficacy and output of the model. 53

4.4 Plot of sEMG data (black line) from the D_f dataset, showing true drill labels (red line), predictions from a model trained with spike loss regularization (blue dashed line), and predictions from a model trained with binary cross-entropy loss (green dashed line). The x-axis represents time, with each index corresponding to a 1 ms interval, consistent with a 1 kHz sampling rate. The y-axis measures the signal amplitude in millivolts. The model utilizing spike loss regularization accurately predicts the sequence of true drill labels, whereas the model trained with binary cross-entropy loss fails to identify the initial drill label window. 59

4.5 Plot of sEMG data (black line) from the D_f dataset, illustrating true drill labels (red line), predictions from a model trained with spike loss regularization (blue dashed line), and predictions from a model trained with binary cross-entropy (BCE) loss (green dashed line). The x-axis represents time, with each index corresponding to a 1 ms interval, consistent with a 1 kHz sampling rate. The y-axis measures the signal amplitude in millivolts. Both models fail to detect the initial drill label subwindow. The spike loss regularization model accurately predicts subsequent drill labels, whereas the BCE loss model incorrectly interprets a subwindow at approximately the 16^{th} second as non-drill activity, resulting in a false negative. 60

5.1 Illustrations of normalised power spectral density (PSD) for sEMG signal, denoted as $x[n]$, and MIA, represented as $q[n]$. These plots are generated based on the sEMG data and recorded noise obtained during the experiments. 63

5.2 Unfolded architecture of the proposed BLSTM-RNN model. The model processes raw sEMG data sequences, with each sEMG sample corresponding to an LSTM block. This ensures that the number of RNN time steps matches the total number of input sEMG samples. Both the forward layer vector $\overrightarrow{h}$ and the backward layer vector $\overleftarrow{h}$ are concatenated after the first and second hidden layers. Each hidden vector at a specific time step contains a feature vector with a length of 128. After concatenation, the resulting feature vectors, now of length 256, are input to a fully connected (FC) layer with a sigmoid activation function. This returns the probability that the corresponding sEMG sample belongs to a MIA. 64

5.3 The illustration of an sEMG recording from the experiments, alongside the true onset/offset labels and predicted labels by various methods: Envelope, SampEn, M-ALED, ACD, and BLSTM RNN. Notably, the high MIA interference observed between the 40^{th} and 60^{th} seconds could be attributed to sensor flapping caused by weakened attachment on the body, a rare occurrence typically following high-intensity exercises. This happens when medical tapes lose their adhesiveness, leading to sensors flapping against the body. The depicted data segment specifically aims to showcase the BLSTM RNN's precise performance under such extreme conditions. 70

5.4 The plot of the relationship between the mean frequency (Hz) of the MIA and the parameter b, as described in (5.12). The parameter b is uniformly selected from a range between 0.1 and 0.5, allowing for the exploration of the impact of varying levels of b on the mean frequency of the MIA. 72

5.5 The illustration of the block diagram for the network structure of the UED model. The model processes an sEMG window of $(1, 128)$, yielding a denoised vector of equivalent dimensions. Within the encoder block, the channel size is augmented by a factor of 4 in the first layer, and by a factor of 2 for subsequent layers, while PL^2 reduces the sequence length to half through length-wise average pooling. Post each Encoder, its output is fused to the corresponding level's decoder output via a copy and crop technique. The decoder block extends the length and reduces the channel size, while P$^2c, L$ performs average pooling by half for both channel and length dimensions. The Con | Att block denotes the concatenation over the channel axis followed by the application of the channel attention module. 75

5.6 The figure outlining the encoder block employed within the main model architecture, as referenced in Figure 5.5. The block receives an input of shape (C,L), representing channel size and sequence length, respectively, and yields an output tensor of shape $(2^{i+1}C,L)$, where i is the layer number. Four parallel convolution blocks are implemented in this model, each with an output filter size of 4 and varying kernel sizes—3, 5, 7, and 9. Here, 'Att' represents the channel attention block, and P_c^4 denotes channel-wise average pooling that reduces the channel size by a factor of four. Post-encoder, the sequence length is halved via average pooling P_L^2. The model, by employing diverse kernel sizes and filter outputs, directs attention to each convolution block to discern the optimal channels. 76

5.7 The figure of a schematic representation of the decoder module, referenced in Figure 5.5. The module receives an input tensor of dimensions (C,L) and outputs a tensor of dimensions $(\frac{C}{4},2L)$. Essentially, the module performs operations inverse to those of the encoder, modifying the shape of the vector correspondingly. Mirroring the actions of the encoder block, this model deploys 4 parallel convolution blocks, succeeded by channel attention and P_c^4 for channel-wise average pooling to size $\frac{C}{8}$. A distinction is that the convolution blocks employ an output filter size that is half of the input channel size. The concatenated vector is then integrated with the convolution transpose block, utilizing a stride and kernel sizes of two, thus increasing the sequence length by 2. The final convolution block further reduces the channel size to $\frac{C}{4}$. 76

5.8 This plot shows the original noisy signal, x, the clean signal with no MIA, $\hat{x}$, and the denoised signal from the UED model, $\bar{x}$. It provides a clear comparison, illustrating how the UED model improves signal SNR by denoising the recorded signal. 79

6.1 The schematic representation of the proposed model. The model accepts two input tensors of shape (B,C,L), corresponding to low frequency-band, x_l, and high frequency-band, x_h. B corresponds to the batch size, C refers to the number of channels, and L indicates the sequence length. Convolution layers are denoted as $\mathbf{f}@Conv\mathbf{k}\times\mathbf{s}$, where f, k, and s represent the output filter size, kernel size, and stride, respectively. The variable N indicates the number of classes in the dataset, with specific values for different datasets: 49 for D_n, 17 for $D_{n,b}$, and 9 for D_f. 85

6.2 The schematic representations of the KD block. The KD block, presented in Chapter 4, has undergone slight modifications: notably, the stride is set to 2 to downsample the signal length, and an initial 1D convolution has an output channel size of 16. As a result, the KD block's channel output is 256, compared to the original 512. The block integrates an initial 1D convolution, parallel D blocks with diverse kernel sizes, a concatenation process with a residual tensor, an attention segment, and a summation with another residual. This transforms the input tensor from a shape of (B,C,L) to $(B,16C,\frac{L}{2})$. 86

6.3 Confusion matrix illustrating the classification performance for 9 classes in the D_f dataset. Correct predictions for each class are indicated by diagonal elements, whereas misclassifications are represented by off-diagonal elements. The color gradient in the matrix highlights the number of instances in each category, effectively demonstrating the model's accuracy and discriminatory capability in classification tasks. 88
6.4 Confusion matrix depicting the classification accuracy over 17 classes in the $D_{n,b}$ dataset. Diagonal elements indicate the correct predictions per class, whereas off-diagonal elements represent misclassifications. The color gradient visually demonstrates the volume of instances in each category, underscoring the model's accuracy and precision in distinguishing between classes. 89
6.5 Illustration of the few-shot metric learning (FSL) framework. The framework accepts a pair of data as input and computes their distance. Function F_θ learns representations that group same-label samples together and separates different-label samples. These representations are then passed to function G_ϕ, which calculates the distance between the paired feature vectors. 92
6.6 Comparative visualizations of the conventional Triplet Loss and the proposed Regulated Centroid-Nested Triplet Loss ($\mathcal{L}_{\text{RCTL}}$). (a) depicts the structure and relationships of the Modified Triplet Loss, underscoring its reliance on large triplet datasets and the importance of triplet selection. Conversely, (b) showcases the distinctive features of the $\mathcal{L}_{\text{RCTL}}$, including centroid regulation and intra-class nested regions, providing a clearer and more insightful perspective on data distributions. In the visualizations, circles filled with the same color belong to the same class (e.g., all black filled circles are one class, and all white filled circles are another). Furthermore, the plus signs at the center of (b) represent the learnt centroids. Collectively, these figures highlight the advancements made to overcome the traditional triplet loss method's constraints. . . 97

7.1 This illustration presents the cumulative recruitment percentages for muscle groups: biceps femoris (BF), semitendinosus (Se), adductor longus (AL), and soleus (So) across two athletes performing the hamstring claw exercise. In plot (a), the athlete exhibits low compensation, consistently showing a right-dominant asymmetry. Conversely, plot (b) showcases an athlete with high compensation, where the reduced activation in the right hamstring is offset by heightened activation in the left hamstring. 105
7.2 This figure delineates the cumulative recruitment percentages of left and right muscle groups—biceps femoris (BF), semitendinosus (Se), adductor longus (AL), and soleus (So)—for two athletes undertaking the hip flexion exercise. In subplot (a), the athlete displays a pronounced bilateral similarity of 99%, indicative of a substantial overlap in the regions. Conversely, subplot (b) illustrates an athlete with a bilateral similarity of 51%, highlighting the disparities in cumulative recruitment between the left and right leg, suggesting the utilization of distinct muscles in each limb. 106

7.3 This figure illustrates the juxtaposition of external load force imbalance percentage (y-axis) against the internal load sEMG recruitment anomaly percentage (x-axis) during the pre-season screening for Manchester United Football Club. Red markers denote players who sustained hamstring muscle injuries within the first three months of the 2022/2023 season, whereas black markers represent players without any injuries during the same timeframe. The graph suggests the potential utility of internal sEMG data and recruitment anomaly metrics as risk indicators in pre-season evaluations, complementing external load metrics. 111
7.4 Block diagram illustrating the deep metric learning framework for anomaly detection. Raw sEMG data serve as the model's input, processed by a twin feature extractor F_θ with shared weights. These feature vectors subsequently inform the distance function G_ϕ, which produces the injury prediction. Importantly, the support set is restricted to drill types consistent with the query, underscoring the significance of identifying the drill class. 115
7.5 A depiction of the centroid (symbolised by a black dot) with corresponding feature vectors for safe (u_s in green), risky (u_r in amber), and injured (u_i in red) states. The distance thresholds for safe (β_1), risky (β_2), and injured (β_3) are also highlighted. This diagram demonstrates how the loss function maintains safe vectors close to the centroid while setting the risky and injured vectors at regulated distances using their respective thresholds. 116

A.1 Illustration of Neurocess sensors on medical-grade adhesive tape prior to attachment to the body. 145
A.2 Illustration demonstrating the attachment of Neurocess sensors on the body. This image is solely for the purpose of showing sensor application and does not represent actual data collection for the D_f dataset. 145
A.3 Illustration showing the Neurocess sensor system on the training ground. 146

List of tables

2.1 Hierarchical classification of acute muscle injuries based on anatomical impact [119] 11

4.1 Layer-by-layer breakdown of the CNN model detailing layer type, input shape, and output shape for each layer. The module processes an input of shape (8,512), representing channel size and sequence length, and extracts a feature vector with a shape of (32,1). 51
4.2 Sizes of train, validation, and test datasets corresponding to different subwindow sizes. 56
4.3 F1 scores achieved by the proposed model when comparing the effects of spike loss and BCE loss across different subwindow sizes. 57

5.1 Optimized λ_1 and λ_2 values per metric. 67
5.2 Overview of training, validation, and test datasets including total windows, samples, duration, and label distributions. 68
5.3 Accuracy, F1 score, sensitivity, specificity, and precision metrics for the state-of-the-art methods and proposed method. 69
5.4 Confusion matrices (in percentages) for onset and offset states during experiments across all the mentioned techniques. The percentages are derived based on the entire sample count, offering a streamlined view of distribution. 69
5.5 Adjustment rates with corresponding mean SNR 73
5.6 Comparison of macro and micro F1 scores at different adjustment rates for datasets D_n and D_f 74
5.7 Comparative analysis of performance metrics across varied depth sizes and attention mechanism for the proposed UED model 78
5.8 Comparative analysis of performance metrics for contemporary methods and the proposed UED model with and without BLSTM RNN detection 79

6.1 Comprehensive summary of preprocessing steps and training/test set sizes for datasets D_n, $D_{n,b}$, and D_f 84
6.2 Comparison of the proposed model's accuracy with other state-of-the-art models across the D_n, $D_{n,b}$, and D_f datasets, evaluated at a 200 ms window size. 87
6.3 Summary of classification metrics for a multiclass prediction model, displaying precision, recall, and F1 score for 9 classes in D_f. 88

6.4 Summary of classification metrics for a multiclass prediction model, displaying precision, recall, and F1 score for 17 classes in $\mathcal{D}_{n,b}$. 89
6.5 Details of the deep learning model acting as G_ϕ, delineating the layer type and output shapes. 95
6.6 Comparison of accuracy (%) among the presented models for 9-way classification using 5, 10, 20, and 50-shot learning. Traditional feed-forward training of the proposed deep learning model in Section 6.1.4 without the FSL technique is included for the 50-shot scenario only. 98
6.7 Average class-based precision, recall, and F1 score for a 50-Shot, 9-Way classification. This table presents a comparative analysis of three different methods: $\mathcal{L}_{\text{RCTL}}$, $\mathcal{L}_{\text{T}}$, and w/o FSL. For each class (1 through 9), the table displays the precision, recall, and F1 score metrics, with the highest F1 score for each class highlighted in bold. 99

7.1 Median Coefficient of Variation for Peak Median Values Across Various Muscle Groups 110
7.2 Distribution of dataset size for safe and risky MD+1 measurements. 118
7.3 Total number of players, and counts of risky and safe measurements per drill 120
7.4 Test set performance metrics for the DML and UGAD models. 121

A.1 Comparison of the sEMG data acquisition systems 144

Chapter 1

Introduction

1.1 Overview and Motivation

Surface electromyography (sEMG) detects electrical signals in muscles during neuromuscular activation. While sEMG finds prevalent application in areas like prosthetic hand control, medical diagnosis, and rehabilitation, its integration within sports science is comparatively limited and largely underexplored. In the context of sports science, sEMG primarily serves to investigate muscle activation patterns and the magnitude of muscular activation during distinct exercises.

In professional football (soccer), injuries not only lead to significant financial burdens but also adversely impact the club's success (see Section 1.2.1). sEMG provides essential insights into muscle physiology, presenting significant potential for injury prevention. Accurate analysis of sEMG data can help prevent sports injuries by assessing muscle injury risk factors. However, a noticeable discrepancy exists between academic research on sEMG and its practical application in sports (see Section 1.5).

There is an abundance of sports science studies employing sEMG [156]. However, there is a lack of studies analyzing sEMG with machine learning techniques in sports science. While some research uses electromyography for sports like running [62, 78, 145], cycling [43, 124], rugby [179], and golf [129, 149] to analyze muscle synergy [36, 17], amplitude [124], timing [103, 117], and frequency content [188, 32] of sEMG, there is still a lack of research combining deep learning techniques with sEMG data for sports science applications.

This research seeks to mitigate muscle-related injuries in athletes by introducing a framework based on sEMG data collected during sports activities (see Section 1.3). Using deep learning, the objective is to refine sEMG analysis and provide a more accurate assessment of athletes' injury risks, bridging the academic-practical gap (see Section 1.4).

Given their high occurrence in football (see Section 1.2.2), this research concentrates on injuries to the hamstring, adductor (often referred to as the groin), and the soleus (a primary muscle in the calf). These injuries are collectively abbreviated as HASI. By applying innovative signal processing techniques and deep learning models, this thesis addresses the prevailing challenges of sEMG applications in the sports domain. The proposed models are applied to real-life data collected from 45 elite athletes from two English Premier League (EPL) football clubs, encompassing 2.5 seasons of continuous recordings. The proposed methodologies are validated using publicly available datasets

and comparisons with state-of-the-art techniques. The aim is to offer solutions and overcome obstacles in accurate sEMG evaluation and injury prevention in professional sports.

1.2 Introduction To The Problem of Interest

1.2.1 The Cost of Injury

This section highlights the impact of muscle injuries in football (soccer). Injuries in football bring two significant setbacks: a financial burden and a negative effect on a team's success. An injury not only deforms the player's and consequently the team's performance [11, 56, 73, 163, 162], but also leads to economic losses [52, 134].

Studies indicate that a player's absence due to injury is strongly connected to a decline in team performance [170, 47, 6] and a decreased chance of success in domestic and European competitions [15, 109]. Research indicates that approximately 136 days of player absence leads to a loss of one league point, while 271 days results in a one-place drop in the league standings [52]. Over an 11-year study of 24 European Champions League Clubs with 7,792 injuries, Hägglund et al. found that teams with fewer injuries ranked higher and earned more points per match [109].

The top 5 European football leagues, namely the English Premier League, La Liga, Serie A, Bundesliga, and Ligue 1, offer an average salary of £1,923,776 per season, equating to £5,271 daily [134]. When a player suffers an injury, these wages often go unutilised. Current evaluations indicate that, besides other health-related expenses, muscle injuries result in a loss of 1,070,714 €/1000 hours in these leagues [134]. Concurrently, Ekstrand et al. report a substantial loss of 500,000 € for each month a player is injured [48]. Complementing this, a recent analysis of injury costs in the EPL between 2012 and 2017 shows that injuries lead to an average annual financial setback of £45 million for an EPL team [52].

1.2.2 Prevalence of Muscle Injuries in Football

Injuries, besides their economic implications and impact on sporting success, occur frequently in football [157, 174, 109]. Muscle injuries, in particular, are prevalent, constituting up to 37% of all time-loss injuries in men's professional football [7, 51, 72, 75]. A study on European professional football indicates that a club with a 25-player squad anticipates around 15 muscle injuries per season, with muscle injuries making up over a quarter of all injury-induced downtime [50].

Another study, covering 26 football clubs and 1,401 players across 10 European countries from 2001 to 2010, reveals that out of the 6,140 recorded injuries, 2,123 (35%) are muscle injuries [74]. A total of 92% of muscle injuries impacted the four primary muscle groups in the lower extremities: hamstrings (37%), adductors (23%), quadriceps (19%), and calf muscles (13%) [50]. Injuries to the adductor, hamstring, and calf muscles occur more frequently during the competitive season, while quadriceps injuries are more prevalent in the preseason [50, 52]. The higher incidence of these injuries during the season aligns with their correlation to increased load and intensity [134, 74]. Consequently, these muscle groups become the primary focus of this research for preventive interventions.

Among all muscle injuries, hamstring injuries stand out due to their frequency, recurrence rate, and associated time-loss [65, 87, 7]. For example, researchers document 723 hamstring injuries in six consecutive EPL seasons, resulting in athletes missing a cumulative total of 21,016 training days [49]. The Football Association's initial audit of injuries indicates that hamstring injuries constitute 12% of all injuries [87]. Given its significant impact on the sport, researchers extensively investigate risk factors associated with hamstring injuries. Green et al. published articles that collectively report on 8,139 hamstring injuries and 967 recurrences in 71,324 athletes, underscoring the extensive research in this area [20].

Despite significant advancements in sports science concerning risk factors, injury frequency remains high in sports [157, 174, 7]. A predominant challenge is the infrequent use of intrinsic feedback tools throughout the season due to their lack of mobility [74]. While tools like isokinetic dynamometers provide valuable intrinsic data, their immobility makes regular in-season assessments logistically challenging. Consequently, clubs often reserve their use mainly for pre-season evaluations. However, many injury risk factors evolve during the season, with pronounced fluctuations in training loads, match rates, and athlete physiology [16, 111]. This underscores the need for continuous, season-long injury risk analysis. Using sEMG data collected consistently, this thesis aims to enhance in-season injury risk assessments, moving beyond the current practice of limited pre-season medical checks. While sEMG offers a promising approach, its use in sports settings has certain challenges, especially in signal analysis. This research aims to utilise sEMG data effectively and address its inherent challenges, with the goal of establishing a robust injury prevention framework for the sports sector.

1.3 Practical Application of Injury Prevention Framework and Data Collection

The thesis focuses primarily on football, specifically aiming to understand and predict prevalent muscle injuries, which are financially burdensome within the sport. More precisely, the investigation relates to the EPL, employing data from two clubs, which spans 2.5 seasons. Detailed insights into the dataset are explored in Chapter 3.

In order to align the research with practical, real-world scenarios, data are consistently collected one day after each match day (MD+1). When players exhibit high-risk markers, an additional data collection session is conducted using sEMG sensors on MD+3. This approach complies with industry norms, as the prevalent practice among football clubs is to conduct risk assessments on MD+1 or MD+2 due to the escalated load and fatigue experienced by players on match day. The study predominantly adheres to the MD+1 data collection protocol since the collaborating football clubs often play matches every three days, validating MD+1 as the optimal time for sEMG-based risk assessment.

Every MD+1, physiotherapists positioned sensors on four bilateral muscle groups, a total of eight sensors. These were strategically placed on the lateral and medial hamstrings, adductors, and soleus. The aim is to predict the likelihood of injuries based on the sEMG data collected during the exercises supervised by club strength and conditioning coaches and physiotherapists. It is vital to acknowledge

that the study strictly evaluates hamstring, adductor, and soleus muscle injuries, excluding all other injury types.

The proposed framework is specifically designed to monitor and adapt to the ever-changing physical demands on an athlete, which fluctuate with the varying intensity levels experienced during matches and training. Central to this framework is the use of sEMG data, gathered the day after a match (MD+1), to predict the risk of hamstring, adductor, and soleus muscle injuries (HASI). This prediction covers the period leading up to and including the next match day. It provides an informed and timely injury risk assessment by incorporating the latest data on the players' physical activities.

In conclusion, this research adopts a practical and realistic approach, utilising actual data from football players and relying on well-established risk assessment techniques. The emphasis is primarily on the prediction and prevention of muscle injuries within football. The methodology is straightforward and aligns closely with the practical requirements of sports science, presenting a promising contribution to more effective injury prevention strategies in the sport.

1.4 The Proposed Injury Prevention Framework

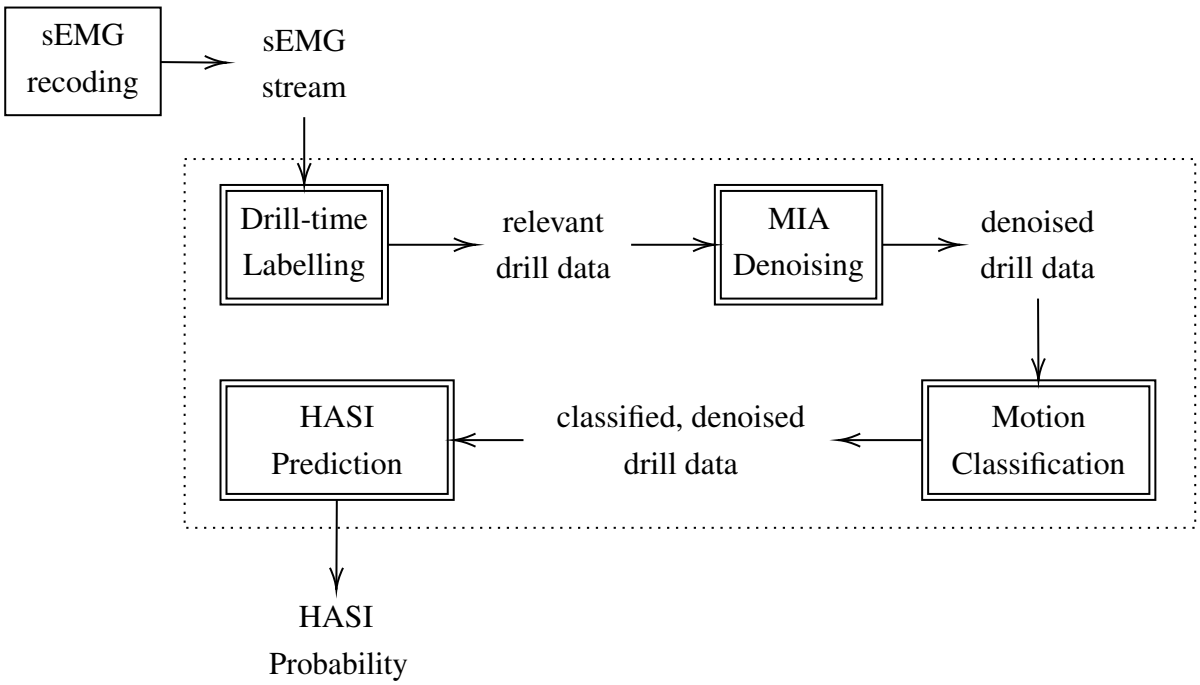


Fig. 1.1 Schematic of the sEMG-based injury prevention framework. The flowchart details the process sequence beginning with sEMG recording, followed by Drill-time Labelling, MIA Denoising, and Motion Classification. The output from Motion Classification is then input into the HASI Prediction model, which calculates the HASI likelihood. The dashed outline denotes the entire framework, which processes the sEMG stream to yield a HASI probability output. Models within the framework are indicated by double-bordered squares.

This research introduces a framework focused on injury prediction within professional sports. The framework is designed to process sEMG data streams recorded from athletes during training sessions

and output the HASI probability. As illustrated in Figure 1.1, the framework is composed of a series of interconnected models, each contributing significant insights for injury prevention. Every step within the framework is strategically linked, with the outcome of one serving as the input for the next.

The framework comprises the following systematically arranged steps:

1. **Drill-time Labelling** (Chapter 4): This initial step focuses on labelling specific times within the recorded sEMG stream that correspond to the moments when the athlete performs a drill. It filters out the relevant data segments associated with the drills from the entire dataset, ensuring that unrelated physical activities or contractions are excluded from subsequent analyses.

2. **Motion-Induced Artefact (MIA) Denoising** (Chapter 5): This crucial step enhances the clarity of sEMG signals by eliminating Motion-Induced Artifacts (MIA), which could otherwise impact the accuracy of subsequent signal processing and machine learning models. It aims to regenerate clean sEMG signals while preserving essential sEMG data, ultimately yielding an output vector of the same size devoid of MIA. Importantly, this step utilises the labelled sEMG segments, specifically the output of the first step, known as drill-time labelling.

3. **Motion Classification** (Chapter 6): By utilising the MIA denoising process outcomes, the motion classification step discerns the specific drill being performed within the sEMG signal segment. This categorisation is critical in the context of injury prevention analysis, as muscle recruitment patterns exhibit variations dependent on the specific drill type. Therefore, determining the drill type before conducting any analysis or estimation becomes imperative. This step automates the classification of drill types.

4. **HASI Risk Estimation** (Chapter 7): This conclusive step calculates the probability of injury for the player by analysing labelled, denoised data segments and the designated drill type. The estimated risk is represented by a value ranging from 0, indicating the least probability, to 1, indicating the utmost probability of injury.

In summary, this framework allows for detailed and clear sEMG analysis, offering reliable results crucial for avoiding HASI in sports.

1.5 Challenges of sEMG Utilisation in Sports Settings

The principal gap between academic sEMG research and its practical applications in real-world sports settings primarily stems from the contrasting environments in which they operate. By their nature, academic settings are controlled environments where every aspect, from data collection to task execution, is meticulously managed. Any extraneous interference is negligible in such settings. Conversely, practical sports settings are characterised by simultaneous interferences and dynamic, uncontrolled data collection conditions. Each step within the framework poses its own set of challenges that are uniquely pertinent to sports settings. The challenges inherent to each step are detailed as follows:

1. **Drill-time Labelling** (Chapter 4):

    - The drill-time labelling step within the proposed injury prevention framework differs from the classical onset sEMG detection task. The traditional task distinguishes contraction, i.e., sEMG signal, from rest, i.e., baseline noise. In sports settings, drills involve repetitive movements. Between drills, athletes may walk, reposition themselves, prepare for the next drill, or rest. Consequently, there may be incidental sEMG activity and contractions in these intervals, which should not be categorised as drills since they are unrelated between-drill contractions. However, such intervals may still encompass sEMG activations as muscles engage during walking or other physical activities. Unlike traditional onset detection, which differentiates muscle activity from inactivity, labelling in this framework identifies specific activities and unrelated motions, like walking between drills.
    - The complexity of the process is elevated due to the incorporation of 8-channel data designed to monitor different muscle groups. Typically, most onset detection methods are limited to a single-channel sEMG stream. In contrast, this thesis encompasses the measurement of eight distinct muscles, encountering scenarios where some muscles are not activated whilst others are, as exemplified in single-leg drills where athletes only engage the muscles in one leg, leaving the other in a state of inactivity. In the context of multichannel sEMG drill labelling, it is imperative for the model to evaluate the muscles collectively, making cohesive determinations on whether the detected activities are components of a drill or not.
    - During drills, athletes might opt for short breaks, especially when they find the exercises difficult. In such intervals, the stream is likely to exhibit an absence of sEMG activation. Nevertheless, to maintain labelling accuracy, it is crucial that these brief resting phases are accurately incorporated within the drill label. This incorporation presents an added layer of complexity to sEMG drill labelling. It is essential to encompass the temporal patterns, meaning that even if athletes take short, intermittent breaks within the drills, such instances should be meticulously considered and labelled as integral parts of the drills.

2. **Motion-Induced Artefact (MIA) Denoising** (Chapter 5):

    - MIA and sEMG signals exhibit overlapping frequency spectra and possess non-stationary characteristics. Consequently, the use of advanced signal processing methods to denoise MIA may inadvertently lead to the elimination of crucial sEMG data.
    - MIA events occur randomly and exhibit no consistent pattern, arising specifically from sudden impacts on sensors or collisions involving the athlete's body. This unpredictability necessitates a targeted approach to maintain data integrity: instead of applying denoising techniques to the entire signal, it is more effective to isolate and treat only those regions of the signal that are affected by MIA. This approach eliminates the need for full sEMG signal stream reconstruction, preserving the integrity of valuable sEMG data, especially

in unaffected segments. However, implementing this strategy is challenging due to the unforeseen timing of MIA occurrences.

3. **Motion Classification** (Chapter 6):

   - In the realm of sports, collecting adequate data for deep learning models is challenging. Given the extensive variety of exercises, accumulating comprehensive data for each is unfeasible [40]. Even when the focus is narrowed to a few pivotal exercises, the inflexible schedules of elite athletes seldom permit thorough data collection. This absence of data imposes significant challenges, necessitating models to adapt to unseen classes with limited available samples, thereby complicating the task of motion classification in sports settings.
   - Hand gestures are characterised by predominantly steady-state, or isometric, contractions, with only brief transient states occurring at the onset [88]. In contrast, lower extremity activities, essential to soccer training, manifest a continuous dynamic sEMG signal, transitioning between eccentric and concentric contractions [96, 145, 5]. This necessitates models to determine temporal sequential differences, whereas models for hand recognition tasks primarily focus on the overall muscle activation rate, often overlooking the temporal patterns.
   - The chosen lower extremity activities in this study exhibit high similarity, with only minor variations observable. For instance, exercises like the glute bridge and elevated glute bridge are nearly identical, distinguished only by slight differences in angle and positioning [108]. This pronounced similarity among motion classes complicates the model's task, making distinctions between them more challenging.

4. **HASI Risk Estimation** (Chapter 7):

   - Every athlete exhibits unique contraction and fatigue patterns, influenced by individual physiological and anatomical differences, medical history—including past surgeries which can impact muscle use—and the distinctive ways they have learnt to execute activities during their development as athletes. These variances necessitate tailored analytical approaches; a model generalized across diverse athlete profiles might generate false positives or negatives in injury estimation due to the inherent disparities in sEMG data among athletes. For instance, sEMG data from an athlete with a history of muscle surgery might be classified as anomalous when compared to the majority, but this could represent their normal baseline. A generalized anomaly detection model might thus yield inaccuracies, which are critical when developing injury prevention models. Accordingly, injury prediction models must accommodate these subjective differences and be meticulously tailored to individual subjects to enhance their predictive accuracy and reliability.
   - Collecting data within professional sports contexts, particularly from professional athletes, presents inherent challenges, with acquiring adequate injury data posing further difficulties. Consequently, models developed in such contexts must be proficient in adapting to anoma-

lies or the athletes' baselines with minimal available samples. This adaptability is crucial to adjust the scarcity of comprehensive datasets in the sports domain.

- The existing literature reveals a notable absence of signal processing-based feature extraction approaches, raising questions about what attributes to consider in a player's sEMG data for injury prediction. While a handful of features are acknowledged, establishing definitive thresholds remains elusive. Given the scarcity of available data and the inherent subject dependency issues, validating such thresholds and formulating features for injury models is a considerable challenge. The absence of identified features signalling the risk of injury and the lack of universally agreed-upon thresholds for such features in the existing studies necessitates independent validation of any proposed technique with the relevant data.

1.6 Contribution and Relevant Publications

The thesis presents novel methods for sEMG data interpretation using deep learning methodologies. One significant contribution of this work is the thorough collection and examination of a distinctive sports dataset, including data from 45 elite athletes over the span of 2.5 seasons from two prominent English Premier League football clubs. Furthermore, the study sheds light on the real-world implications of sEMG within a sporting environment, addressing an existing gap in the sports industry. By introducing a robust framework, this thesis enhances the relevance of sEMG technology, marking its significance in the football realm as a crucial data source. The main contributions of this thesis are detailed as follows.

- This research introduces a comprehensive dataset, encompassing regular sEMG recordings from 45 players across 2.5 seasons from two EPL clubs. The dataset, rich in muscle injury records, is curated following a specific protocol that aligns with the clubs' requirements and constraints, focusing on data collection one day post-match day (MD+1) for practicality. This represents a unique effort in collecting elite athlete data while considering industry standards and practical demands.

- The thesis proposes a deep learning model designed to pinpoint the exact time that athletes undertake drills given the whole stream of recording. By processing 8-channel signals, the model discerns temporal patterns characterising sEMG-based drills and the inter-muscle activation sequences. Given the periodic activations in real-world sports settings, this model is proficient at distinguishing genuine drill-related contractions from non-specific ones.

- An approach targeting motion-induced artefact (MIA) cancellation is outlined. This system first identifies MIA signals, then utilises a deep learning model to denoise affected sEMG sections.

- A motion classification model is presented, with performance benchmarking against contemporary techniques using a publicly available hand gesture dataset. Beyond its superior performance

with this dataset, the model demonstrates robust accuracy in classifying lower extremity drills within the football dataset.

- Recognising the constraints of data collection in sports environments, the research introduces a deep metric learning strategy. This approach leverages publicly available datasets for training, subsequently adapting to the specific demands of football motion classification.
- Essential sEMG signal features are introduced, which can be employed for various physiological insights in risk assessment without requiring a deep learning-based feature learning scheme.
- A pioneering injury prediction model is presented, anchored in the hypothesis that altered muscle recruitment patterns indicate injury proximity. Through a unique metric, this model calculates self-distance of players, successfully forecasting injury risk.
- The presented framework comprises distinct modules, each contributing uniquely to sEMG data analysis. Collectively, they formulate an integrated solution with profound implications for injury prevention. This underscores the combined potential of sEMG and deep learning in sports science.

Journals

1. M. Ergeneci, D. Carter and P. Kosmas, "sEMG Onset Detection via Bidirectional Recurrent Neural Networks With Applications to Sports Science," in IEEE Sensors Journal, vol. 22, no. 19, pp. 18751-18761, 2022, doi: 10.1109/JSEN.2022.3198882. (Chapter 5)
2. M. Ergeneci, E. Bayram, D. Binningsley, D. Carter, and P. Kosmas, "Attention-Enhanced Frequency-Split Convolution Block for sEMG Motion Classification: Experiments on Premier League and Ninapro Datasets," in IEEE Sensors Journal, vol. 24, no. 4, pp. 4821-4830, 2024, doi: 10.1109/JSEN.2023.3345731. (Chapter 6)
3. M. Ergeneci, D. Binningsley, D. Carter, and P. Kosmas, "sEMG-Based Deep Metric Learning With Regulated Centroid-Nested Triplet Loss: From Hand Gestures to Elite Soccer Drills in the English Premier League," in IEEE Sensors Journal, vol. 24, no. 5, pp. 6564-6572, 2024, doi: 10.1109/JSEN.2024.3350237. (Chapter 6)
4. M. Ergeneci, E. Bayram, D. Binningsley, D. Carter, and P. Kosmas, "A Novel Framework for Motion-Induced Artefact Cancellation in sEMG: Evaluation on English Premier League and Ninapro Datasets," in IEEE Sensors Journal, accepted May 30, 2024, to be published, doi: 10.1109/JSEN.2024.3404566. (Chapter 5)
5. M. Ergeneci, D. Binningsley, D. Carter, and P. Kosmas, "A Spike-Regularised Deep Learning Approach for 8-Channel sEMG Data Labeling: Hamstring, Groin, and Soleus Drills in English Premier League Players," submitted to IEEE Sensors Journal, minor revision received May 3, 2024. [Manuscript under review] (Chapter 4)

Conference

1. M. Ergeneci, D. Binningsley, R. Saddler, E. Bayram, P. Kosmas, "The cooperation of isometric force test and EMG for hamstring injury prevention" in Isokinetic Conference FOOTBALL MEDICINE – The Pursuit of Excellence, London, United Kingdom, May 27-29 2023

1.7 Thesis Structure

The subsequent structure of the thesis is outlined as follows. Chapter 2 provides an in-depth examination of the background, encompassing discussions on muscle injuries, their classifications, and underlying mechanisms. Furthermore, it provides an overview of the sEMG signal, containing its generation, the key factors influencing it, and essential signal processing techniques. This chapter also presents an overview on deep learning, specifically underscoring the foundational concepts that are integral to the research. Chapter 3 elaborates on the datasets employed in this research, with a focal point being the football dataset collected from 45 elite athletes from two distinguished English Premier League clubs. The chapter outlines the data collection process, the specific muscles analyzed, the drills selected, and the prudent standards maintained during data collection. Chapter 4 introduces the preliminary building block of the proposed framework, clarifying the algorithmic strategies to autonomously label the drill-timing of players from extended sEMG sequences. Chapter 5 delves into the subsequent module, addressing the task of identifying and denoising the motion-induced artefacts in sEMG signals. In Chapter 6, the focus shifts to the framework's third component: motion classification. The chapter compares the proposed methods with established benchmarks using well-known hand gesture datasets. It also considers the challenges of motion classification in scenarios with constrained data availability, proposing novel solutions via deep metric learning. Chapter 7 unveils the concluding segment of the framework, discussing the methodologies formulated to ascertain the potential injury risks in athletes, predicated on their collected MD+1 data. Chapter 8 consolidates the research findings, offers concluding thoughts, and identifies potential directions for future studies.

Chapter 2

Background

2.1 Muscle Injuries: Classification and Mechanisms

Table 2.1 Hierarchical classification of acute muscle injuries based on anatomical impact [119]

A. Indirect muscle injury		
Functional Injury	Type 1: Overexertion-related	Type 1A: Fatigue-induced
		Type 1B: Delayed-onset muscle soreness
	Type 2: Neuromuscular	Type 2A: Spine-related neuromuscular
		Type 2B: Muscle-related neuromuscular
Structural Injury	Type 3: Partial muscle tear	Type 3A: Minor partial tear
		Type 3B: Moderate partial tear
	Type 4: (Sub)total tear	Subtotal or complete tear
		Tendinous avulsion
B. Direct muscle injury		
	Contusion	
	Laceration	

As outlined in Table 2.1, muscle injuries fall into two primary categories: direct and indirect injuries [119].

Direct muscle injuries result from an external force directly impacting the muscle [13]. Examples include contusions, lacerations, or blunt force trauma [119]. These injuries frequently cause disruption of muscle fibres, blood vessels, and connective tissues, leading to hematoma formation, inflammation, and potential muscle fibre necrosis [50]. Due to their unpredictable nature stemming from external sources, these injuries cannot be anticipated using intrinsic physiological risk factors [119].

Conversely, indirect muscle injuries occur intrinsically, often during muscle contraction or overstretching, without a direct external force impact [50]. The usual causes include overexertion, fatigue, or improper biomechanics, primarily affecting the musculotendinous junction [119]. These injuries can vary in severity, ranging from minor muscle fibre tears to complete ruptures [119].

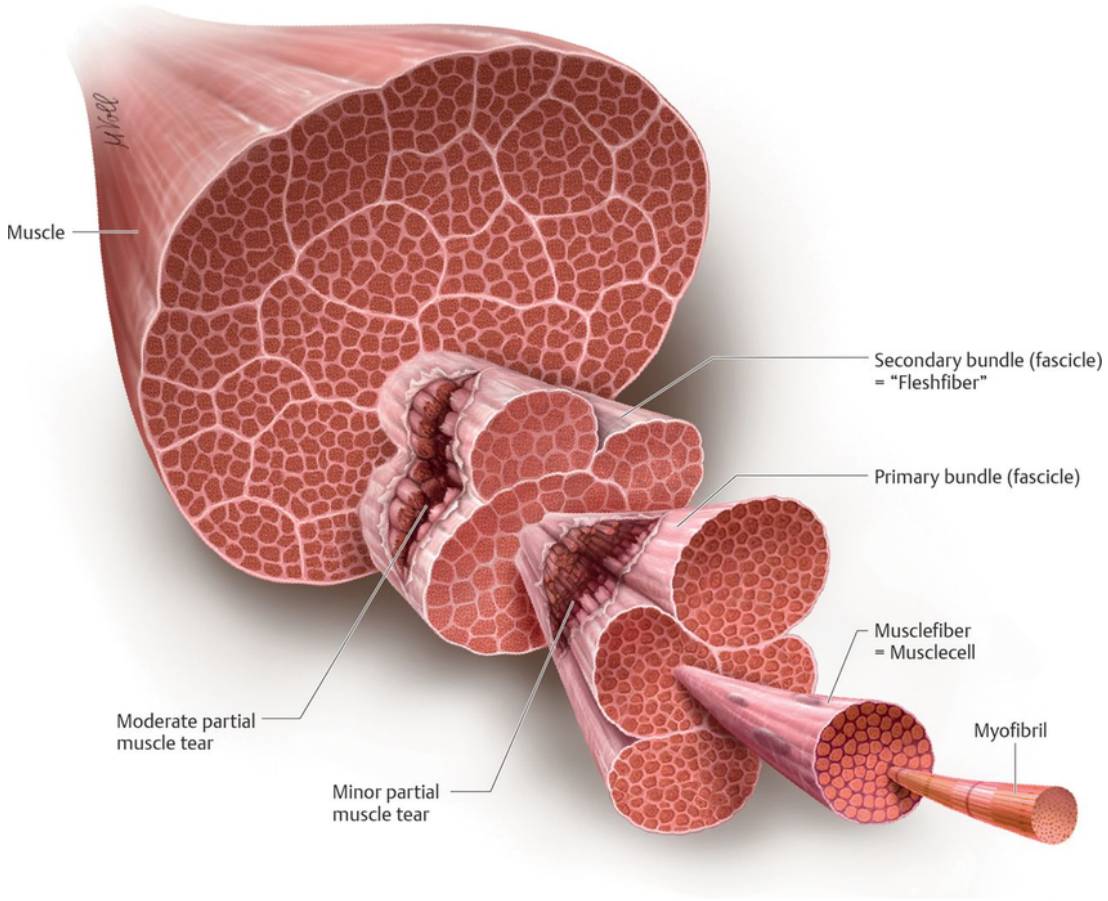


Fig. 2.1 Illustration of minor and moderate muscle tears on a muscle diagram, sourced from [119].

Indirect injuries can be further categorised into structural and functional injuries. While functional injuries result in muscle pain and stiffness due to minor fatigue-induced or neurogenic origins, structural injuries encompass muscle tears or tendon ruptures [119].

Type 1A injuries, known as fatigue-induced muscle disorders, are mainly a result of muscle overexertion without evident structural damage [169, 125]. The root causes for these disorders are often acute overload and muscular imbalance [169, 125, 172]. In contrast, Type 1B injuries or Delayed-onset muscle soreness (DOMS) stem from unfamiliar or intense exercises that cause muscle fibre microtrauma [147]. Players suffering from Type 1 injuries are typically sidelined for 2 to 5 days [119].

Notably, neuromuscular injuries (Type 2) differ from fatigue-induced ones [119]. They relate to dysfunctions in the nervous system and its interaction with muscles, such as issues at the neuromuscular junction or spinal nerve impingement [128]. These injuries, more concerned with the nervous system's inefficiency in communicating with muscles than muscle fatigue, are less frequent compared to Types 1 and 3 [119]. Their incidence is not linked to training and effort dynamics during a season. In our data collection spanning three years across 45 athletes, only one Type 2 injury incident was recorded.

Structural injuries damage the musculoskeletal system's anatomical components, compromising tissue integrity [50, 119, 89, 169, 29, 64, 84]. Type 3A injuries, signifying minor partial muscle tears, often result from forceful or overstretching movements, with a recovery period of 10 to 23 days [71]. Type 3B injuries, moderate partial muscle tears, may need 14 to 26 days for recovery [71]. Lastly,

Type 4 injuries, including (Sub)total tears, require varying recovery durations—up to 70 days for non-surgical treatments and 120 days post-surgery [71].

This research primarily considers Types 1, 3, and 4 injuries attributed to high training loads, match intensity, and accumulated fatigue. These can be monitored through intrinsic risk factor shifts. The study aims to predict fatigue-related injuries, where sEMG is instrumental in tracking muscle

2.1.1 Hamstring: Anatomy and Injury Mechanisms

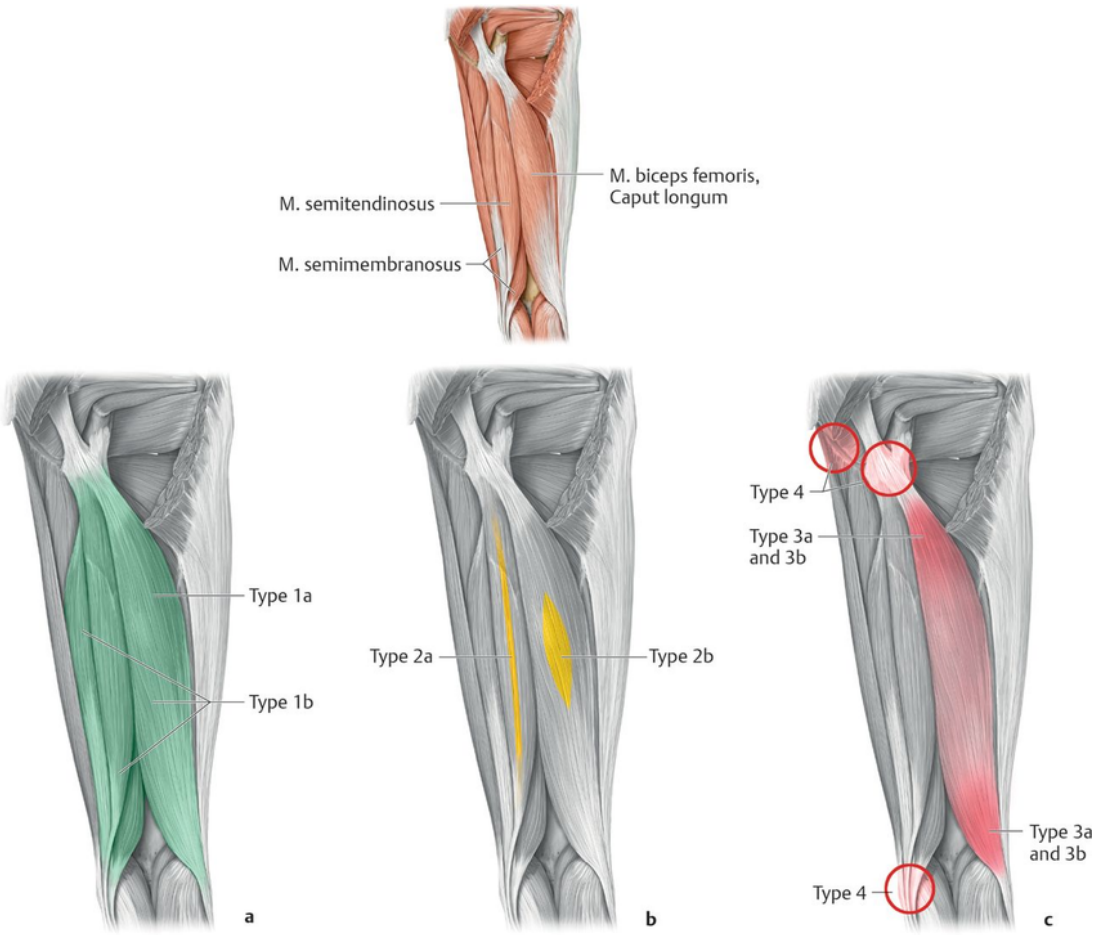


Fig. 2.2 Illustration of hamstring muscle groups and associated injury types, sourced from [119].

The hamstring, an essential muscle group in the posterior thigh, comprises three main muscles: the biceps femoris, semitendinosus, and semimembranosus [30, 123]. Crossing both the hip and knee joints, the hamstring is instrumental in hip extension and knee flexion—key movements in many sporting activities. Notably, the biceps femoris is positioned laterally in the thigh, whereas the semitendinosus is medially situated [144, 121]. Although both muscles contribute to the primary functions of hip extension and knee flexion, they also have distinct roles: the biceps femoris facilitates lateral rotation of a flexed knee, while the semitendinosus supports medial rotation [144]. Explosive and sudden physical activities requiring knee-bending, such as jumping and running, demand substantial hamstring engagement, contrasting with activities of lower intensity like standing or walking [123, 85]. Owing to

their involvement in these core movements and their role in stabilising the pelvic and knee joints, these muscles are vital for athletic performance [85].

Epidemiological studies predominantly demonstrate that hamstring injuries often occur during eccentric contractions, where muscle fibres extend under physical load [22, 169, 158, 132, 85]. These studies indicate that certain eccentric hamstring contractions lead to non-uniform sarcomere lengthening, creating micro-level musculotendinous damage [69, 45]. Repetitive eccentric contractions resulting in such microscopic damages compromise the muscle's integrity, leading to soft tissue tears of varying severity [143]. Figure 2.2 illustrates the injury types discussed in Section 2.1 on each muscle of the hamstrings.

2.1.2 Adductor: Anatomy and Injury Mechanisms

The adductors, located in the medial part of the thigh, consist of several key muscles, including the adductor magnus, adductor longus, adductor brevis, gracilis, and pectineus [99, 61, 118]. Extending from the pelvis to the femur, the primary role of the adductors is hip adduction, pulling the leg towards the body's centre, a fundamental action in many sports [148, 61]. The adductor longus and brevis are positioned anteriorly, while the adductor magnus is situated posteriorly [99]. Notably, the adductor longus is the most commonly injured among these muscles [99]. Apart from hip adduction, these muscles also participate in other hip movements such as flexion, extension, and external rotation [99, 148]. Activities that involve swift lateral motions, like football, significantly engage the adductors [148, 99].

Adductor injuries result from several activities: 35% relate to changes of direction, 29% to kicking, 24% to reaching, and 12% to jumping [148]. Among these, injuries stemming from changes of direction and reaching constitute 59% and fall under closed chain movements, typically involving hip extension, abduction, and external rotation [148]. In contrast, injuries from kicking and jumping, representing the remaining 41%, are categorised as open chain movements. These injuries entail shifts from hip extension to flexion, hip abduction to adduction, and external rotation [148].

2.1.3 Soleus: Anatomy and Injury Mechanisms

The soleus, a calf muscle component located in the lower leg's posterior part, collaborates closely with the gastrocnemius [2]. Originating from the tibia and fibula, it joins the Achilles tendon as its insertion point. The primary function of the soleus is plantarflexion, which involves pointing the toes downward or elevating the heel [80]. This action is most evident during the propulsive phase of gait when the foot pushes against the ground to move the body forward. Within the realm of sports, soleus injuries show a marked prevalence in football, presenting an incidence approximately five times greater than that in tennis, track-and-field, basketball, field hockey, and triathlon [131]. Activities that pose as risk factors for soleus injuries include high-speed running, prolonged running durations, repetitive accelerations and decelerations, and situations demanding performance under fatigue [37, 120, 141, 70, 31].

2.2 Surface Electromyography (sEMG): An Overview

Surface electromyography (sEMG) represents the cumulative electrical activity of active motor units (MUs) detected by electrodes on the skin above the muscle [12, 38, 53]. Physiological changes influence this activity in the muscle fibre membranes. Important characteristics of sEMG, such as amplitude and power spectrum, depend on the properties of the muscle fibres and the timing of MU action potentials [57]. Thus, sEMG captures the peripheral and central aspects of the neuromuscular system [59].

2.2.1 Mechanisms of sEMG Signal Generation

Surface electromyography (sEMG) reflects the aggregated electrical activity of active motor units (MUs) captured by electrodes positioned over the skin covering the muscle [102]. A MU, representing the most basic functional unit for muscle contraction, consists of a single motor neuron and all the muscle fibres it stimulates [102]. Upon activation of a motor neuron, every affiliated muscle fibre within that motor unit contracts simultaneously. The specific sites where a MU stimulates a muscle fibre are termed motor endplates, and the collective area of these stimulation points is referred to as the innervation zone (IZ) [53, 102, 59].

The muscle fibre is surrounded by a cell membrane known as the sarcolemma. Embedded within this membrane are three critical gate structures: the voltage-gated sodium (Na^+) channel, the voltage-gated potassium (K^+) channel, and the K^+/Na^+ ATPase pump [102, 171, 35]. In its resting state, the intracellular environment has a negative charge, typically between -70 to -90 mV, and is enriched with K^+, whilst the extracellular environment is Na^+ rich. This balance is maintained by the K^+/Na^+ ATPase pump. Following neural stimulation at the neuromuscular junction, the voltage-gated Na^+ channel opens, allowing Na^+ ions to rush into the cell. This influx raises the membrane voltage to approximately 30 mV, a process termed depolarisation. Subsequently, as the cell interior becomes positively charged, the voltage-gated K^+ channel opens, causing K^+ ions to exit the cell and reduce the voltage to its resting negative values, known as repolarisation. Upon completion of repolarisation, the K^+/Na^+ ATPase pump activates to restore the K^+ and Na^+ balance, using ATP as its energy source. This pump imports two K^+ ions into the cell while exporting three Na^+ ions. The electric waveform resulting from the depolarisation and repolarisation events is termed the action potential (AP). The sEMG signal fundamentally derives from these APs occurring at the muscle fibre membrane. The region undergoing depolarisation typically spans an area of 1-3mm^2 [102]. After the initial excitation, this zone propagates along the muscle fibre at velocities ranging between 2-6m/s [102, 171]. This excitation wave commences at the IZ and extends bilaterally to the tendon extremities [171].

The muscle action potential propagates across the entire surface of the muscle cell and delves into the cell through specific channels named T-tubules [102]. These T-tubules are closely related to the sarcoplasmic reticulum, a structure responsible for storing calcium ions. As the electrical signal travels through the T-tubules, the sarcoplasmic reticulum releases calcium ions into the muscle cell's cytoplasm.

This influx of calcium is crucial for muscle contraction. Muscle cells contain fundamental units known as sarcomeres, representing the primary sites for contraction. Sarcomeres comprise thin (actin) and thick (myosin) protein filaments. In the presence of calcium, these filaments bind, enabling the myosin to exert force on the actin, causing the sarcomere to contract. When numerous sarcomeres in a muscle cell contract at the same time, the entire muscle fibre contracts. This collective contraction of multiple fibres in the muscle results in the overall contraction of the muscle, facilitating movement or force generation.

Traditional sEMG sensors employ bipolar electrodes, typically spaced 10mm apart, in conjunction with a differential amplifier to measure the voltage difference across two points on the muscle [102]. Utilising differentiation primarily aims to minimise common mode noise, such as the 50Hz power line interference [102, 113]. Given that a MU comprises a neuron linked to multiple muscle fibres, a stimulus induces several APs. When these APs overlap, they form the motor unit action potential (MUAP). Since a muscle contains numerous MUs, sEMG essentially records the cumulative effect of various MUAPs during a muscle contraction [171, 102].

In addition to peripheral factors, central properties of motor units also significantly influence the generation of muscle force and the resulting sEMG signal. One critical central property is the motor unit discharge rate. The discharge rate, or firing rate, refers to the frequency at which a motor neuron sends action potentials to its associated muscle fibers. Higher discharge rates lead to more forceful muscle contractions as more frequent action potentials cause more sustained and stronger contractions of the muscle fibers [38].

Another essential central property is motor unit recruitment. Motor unit recruitment follows the size principle, where smaller motor units are activated first at lower force levels, and larger motor units are recruited as more force is needed [77, 38]. The recruitment and derecruitment of motor units are crucial for fine-tuning muscle force output and play a significant role in shaping the sEMG signal. The combination of varying discharge rates and the sequential recruitment of motor units creates a complex and dynamic pattern of electrical activity, which is captured in the sEMG signal.

Furthermore, motor unit synchronization, where multiple motor units fire simultaneously or near-simultaneously, can enhance the force produced by the muscle and influence the characteristics of the sEMG signal. Synchronized firing can lead to more pronounced and detectable signals, particularly during high-force contractions or rapid movements [180].

By considering both peripheral and central mechanisms, we gain a comprehensive understanding of how sEMG signals are generated. This dual perspective allows for a more accurate interpretation of sEMG data, which is essential for applications in biomechanics, neuromuscular research, and clinical diagnostics.

2.2.2 Characteristics and Factors Influencing the sEMG Signal

The sEMG signal exhibits both stochastic and non-stationary characteristics [57, 59, 113]. Its stochastic nature arises from the inherent randomness of neuromuscular electrical activity, primarily due to the asynchronous activation of multiple MUs [102]. Each MU produces a distinct motor unit action potential (MUAP), and the superposition of these diverse MUAPs contributes to the unpredictable

waveform of the sEMG signal [35]. On the other hand, the non-stationary aspect of sEMG can be attributed to changing statistical properties over time [113, 59]. Factors such as muscle fatigue, variation in motor unit recruitment, and alterations in muscle fibre conduction velocity can lead to fluctuations in the signal's mean and variance, rendering its statistical properties inconsistent across its duration [113, 59].

The amplitude of the sEMG signal, possessing a zero mean, typically spans from microvolts to several millivolts peak-to-peak [140]. The sEMG signal's frequency spans 10 – 400 Hz [113]. Its mean frequency is between 70 and 130 Hz, and its median frequency falls within 80 – 120 Hz [105, 106, 113].

The characteristics and quality of the sEMG signal are influenced by a variety of factors. Physiologically, the variations in the number and type of activated MUs, changes in muscle fibre conduction velocity due to fatigue or pathologies, manifestations of muscle fatigue from prolonged activity, the inherent muscle architecture, muscle temperature, and the depth of muscle from the skin surface play pivotal roles [113, 59, 57]. Technically, factors such as electrode placement relative to the innervation zone, inter-electrode distance, alignment with muscle fibres, the choice between gel-based and dry electrodes, skin conditions including cleanliness and subcutaneous fat layers, potential external electrical interferences, and amplifier characteristics significantly affect the recorded signals [154, 181, 102, 182, 4, 112, 59]. Externally, movement-induced electrode displacement, ambient noise from surrounding electrical equipment, and limb positioning can introduce variability [102, 4].

In the case of this thesis, the data collected included elite players performing sports drills under varying fatigue states, which adds a certain variability to the sEMG data captured. Fatigue causes the athletes to alter the recruitment pattern of muscles; for example, instead of the biceps femoris, an athlete may tend to contract the adductor longus of the same leg more to support the external force and to compensate for the sEMG activity reduction in the hamstring caused by muscle fatigue due to overexertion. Fatigue also affects the amplitude and frequency components of the muscles, as it pushes the model spectral distribution to lower frequency levels due to the decrease in muscle conductivity. This change is crucial in sports applications where performance and recovery are monitored through sEMG signals [115, 68].

2.2.3 Fundamental Approaches to sEMG Signal Analysis

sEMG signal analysis involves processing in the time domain, frequency domain, and time-frequency domain to extract various features. This section elucidates the primary processing techniques in these domains. sEMG signals typically present as lengthy sequential recordings. For effective analysis, researchers often segment these recordings into smaller, equal-length parts, commonly employing the sliding window technique. Given a discrete signal $x[n]$, this technique segments the signal into consecutive, potentially overlapping windows, making it suitable for contexts like time-series analysis. Let W denote the window size, which is the number of samples in each segment, and O represent the overlap between consecutive windows. Thus, each new window includes $W - O$ new samples. The k^{th} windowed segment, $s_k[n]$, is defined as

$$s_k[n] = x[n + k \times (W - O)], \tag{2.1}$$

where, n ranges from 0 to $W-1$, and k is the window index.

Time Domain Features: Time-domain features directly analyze the amplitude and shape variations of sEMG signals over time. They provide a straightforward interpretation of the signal's behavior by computing statistics like mean, variance, or specific patterns from the raw data segments. The key time domain features are provided as follows.

- **Mean Absolute Value (MAV)** represents the average rectified amplitude of the sEMG signal. For the window s_k, MAV is calculated as

$$\text{MAV}(s_k) = \frac{1}{W} \sum_{n=0}^{W-1} |s_k[n]|. \tag{2.2}$$

- **Root Mean Square (RMS)** calculates the square root of the average of the squared values of a signal, emphasizing a sensitivity to larger values and spikes within the signal. For the window s_k, RMS is given by

$$\text{RMS}(s_k) = \sqrt{\frac{1}{W} \sum_{n=0}^{W-1} s_k[n]^2}. \tag{2.3}$$

- **Waveform Length (WL)** quantifies the cumulative waveform length over a window, reflecting the signal's complexity and frequency content. For the window s_k, WL is given by

$$WL(s_k) = \sum_{n=1}^{W-1} |s_k[n+1] - s_k[n]|. \tag{2.4}$$

- **Zero Crossing Rate (ZCR)** measures the rate at which a signal changes sign, indicating its frequency content. For the window s_k, ZCR is determined by

$$ZCR(s_k) = \frac{1}{N-1} \sum_{n=1}^{N-1} |\sigma(s_k[n]) - \sigma(s_k[n+1])|, \tag{2.5}$$

  where the function $\sigma(\cdot)$ is

$$\sigma(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x \leq 0 \end{cases}. \tag{2.6}$$

Frequency Domain Features: Frequency-domain features convert sEMG signals from the time domain to the frequency spectrum, capturing the rhythmic patterns and oscillations of muscle activations. They highlight the signal's frequency content and identify dominant frequency bands and power levels.

The Fourier Transform is crucial in this conversion. It provides a view into muscle activation patterns by changing the sEMG signals into their frequency-domain. Notably, the Power Spectral

Density (PSD) indicates the power distribution over the signal's frequency components. The Discrete Fourier Transform (DFT) for a windowed sEMG signal $s_k[n]$ is

$$S_k[f] = \sum_{n=0}^{W-1} s_k[n] \cdot e^{-j2\pi \frac{fn}{W}}. \tag{2.7}$$

and the PSD is given by

$$P_k[f] = |S_k[f]|^2. \tag{2.8}$$

By computing statistics such as the PSD, mean frequency, and median frequency, frequency-domain features can offer insights into aspects like muscle fatigue, the type of muscle fiber being activated, and the overall spectral properties of the signal. The principal frequency domain features are further detailed below.

- **Mean Frequency (MNF)** characterizes the central tendency of the frequency components in the sEMG signal. For the power spectral density $P_k[f]$ derived from the window s_k, MNF is determined as

$$MNF_k = \frac{\sum_{f=0}^{F} f \cdot P_k[f]}{\sum_{f=0}^{F} P_k[f]}. \tag{2.9}$$

- **Median Frequency (MDF)** identifies the frequency at which half of the sEMG signal's total power is contained, providing a robust representation of the signal's central frequency. For the power spectral density $P_k[f]$ of the window s_k, the cumulative power spectrum $C(f)$ is defined as

$$C(f) = \sum_{\alpha=0}^{f} P_k[\alpha]. \tag{2.10}$$

  Subsequently, the frequency f_{med} at which this cumulative power becomes half of the total power is recognized as the MDF, and it can be represented as

$$MDF_k : C(f_{\text{med}}) = 0.5 \times C(F). \tag{2.11}$$

Time-Frequency Domain Features: Time-frequency domain features analyse the frequency components of sEMG signals over time. These features offer a unified approach, combining insights from both time-domain and frequency-domain analyses. By examining the signal's behaviour in both time and frequency simultaneously, these features highlight patterns such as transient oscillations and localised frequency shifts. These patterns might be overlooked when studying the domains independently. The main time-frequency domain features are described as follows.

- **Continuous Wavelet Transform (CWT)** analyses the frequency components of the windowed sEMG signal s_k across various scales. Utilising a series of scaled and shifted wavelets, the CWT reveals localised time-frequency characteristics. For the windowed signal $s_k[n]$, the CWT in

relation to a mother wavelet $\psi(t)$ is formulated as

$$CWT[a,b] = \frac{1}{\sqrt{a}} \sum_{n=0}^{W-1} s_k[n] \psi \left[\frac{n-b}{a} \right], \tag{2.12}$$

where $\psi \left[\frac{n-b}{a} \right]$ denotes the mother wavelet function which is scaled by a and translated by b. In sEMG analyses, the Mexican hat wavelet is frequently utilised due to its suitability for such signals. It is essentially the second derivative of the Gaussian function, and is given by

$$\psi[n] = \frac{2}{\sqrt{3\sigma}\pi^{1/4}} \left(1 - \left(\frac{n}{\sigma} \right)^2 \right) e^{-\frac{n^2}{2\sigma^2}}. \tag{2.13}$$

- **Discrete Wavelet Transform (DWT)** decomposes the windowed sEMG signal s_k into a hierarchical series of approximation and detail coefficients. This process effectively captures both the high-frequency and low-frequency components of the signal at various resolutions. For the windowed signal $s_k[n]$, the DWT coefficients can be recursively computed using filter banks comprising of low-pass and high-pass filters followed by down-sampling. The coefficients are obtained as:

$$cA[k] = \sum_{n=0}^{W-1} s_k[n] \phi \left[n - 2k \right], \tag{2.14}$$

$$cD[k] = \sum_{n=0}^{W-1} s_k[n] \psi \left[n - 2k \right], \tag{2.15}$$

where $cA[k]$ and $cD[k]$ are the approximation and detail coefficients respectively, and $\phi(t)$ and $\psi(t)$ are the scaling and wavelet functions.

The choice of the wavelet function $\psi(t)$ and the scaling function $\phi(t)$ is crucial, and there are several families of wavelets used in applications, with the Daubechies wavelets being particularly popular in many sEMG analyses.

- **Discrete Short-Time Fourier Transform (DSTFT)** offers a time-frequency analysis of the windowed sEMG signal s_k. This transform divides the signal into overlapping, windowed segments and then computes the Fourier transform for each segment, thus capturing the evolution of the signal's frequency components over time. The application of a window function $w[n]$ is essential as it mitigates the spectral leakage and provides a trade-off between time and frequency resolution. For the windowed signal $s_k[n]$, the DSTFT is represented as

$$DSTFT[\tau, f] = \sum_{n=0}^{W-1} s_k[n] w[n-\tau] e^{-j2\pi fn}, \tag{2.16}$$

where τ denotes the time shift of the window and f signifies the frequency.

In sEMG analyses, the Hann window (or Hanning window) is frequently employed due to its ability to reduce the side lobe leakage. The discrete-time version of the Hann window is given by

$$w[n] = 0.5\left(1 - \cos\left(\frac{2\pi n}{W-1}\right)\right), \tag{2.17}$$

for $0 \leq n \leq W-1$, where W is the window length.

- **Hilbert-Huang Transform (HHT)** is a novel approach designed specifically for non-linear and non-stationary signal analysis, making it suitable for sEMG signals which exhibit such behaviours. HHT comprises two primary steps: the Empirical Mode Decomposition (EMD) and the Hilbert Spectral Analysis (HSA).

  The EMD method decomposes a signal into a collection of intrinsic mode functions (IMFs). An IMF is a function with the same number of extrema and zero-crossings, with its envelopes being symmetric with respect to zero. For a given windowed signal $s_k[n]$, the decomposition process can be represented as:

$$s_k[n] = \sum_{i=1}^{N} c_i[n] + r[n], \tag{2.18}$$

  where $c_i[n]$ are the IMFs, $r[n]$ is the residue, and N is the total number of IMFs extracted from the signal.

  Following EMD, the HSA is employed to derive the instantaneous frequency and amplitude by applying the Hilbert transform to each IMF. The instantaneous frequency $f_i[n]$ and amplitude $a_i[n]$ for the i-th IMF are given by:

$$a_i[n] = \sqrt{c_i^2[n] + \hat{c}_i^2[n]} \tag{2.19}$$

$$f_i[n] = \frac{1}{2\pi}\frac{d\theta_i[n]}{dn}, \tag{2.20}$$

  where $\hat{c}_i[n]$ is the Hilbert transform of $c_i[n]$ and $\theta_i[n]$ is the phase function.

  HHT provides a time-frequency-energy distribution, making it a powerful tool for the detailed analysis of sEMG signals.

2.3 Foundations of Artificial Neural Networks: An Overview

Deep learning, a distinguished subset of machine learning (ML), operates using neural networks of multiple layers, often termed as deep networks. The strength of deep learning lies in its ability to process a wide variety of data types, with its hierarchical network structure playing a key role. As data progresses through each layer of a deep network, specific features are identified and refined, a process known as feature learning. Traditional ML techniques, such as linear regressions or decision trees, rely heavily on manual feature extraction tailored for particular applications. In contrast, autonomous feature learning in deep learning models enables them to discern intricate patterns without explicit

extraction for each specific task. Recent studies of sEMG analyses show that feature learning allows deep learning models to outperform traditional techniques [8, 34, 175, 136].

Additionally, the performance and accuracy of deep learning models improve with larger datasets, highlighting their scalability. However, deep learning presents certain challenges. For instance, intricate models require robust computational capabilities, often necessitating high-performance GPUs. A substantial amount of labelled data is essential for training, and such extensive datasets might not always be available. Furthermore, a notable drawback is the interpretability of deep learning models. Many of these models, especially the deeper architectures, function as 'black boxes', making them less transparent and more difficult to interpret than some traditional ML models.

This thesis employs deep learning (DL) techniques to address the challenges presented in Section 1.5. The ultimate aim is to devise a comprehensive framework for HASI prevention in football. This section introduces the fundamental principles of ANN, spanning its basic components to the performance metrics utilised for validation in this research. The subsequent methods and techniques form the core of the DL approaches adopted in this thesis.

2.3.1 Understanding Deep Learning Processes: Training, Inference, and Testing

Deep learning models undergo three primary phases: training, inference, and testing. During training, a model iteratively adjusts its weights and biases based on the data it encounters. The aim is to minimise the discrepancies captured by a loss function, which measures the differences between the model's predictions and the actual data labels. Upon the completion of training, the model transitions to the inference phase. In this stage, it employs its learned parameters to predict outcomes for previously unseen data. The subsequent testing phase evaluates the model's predictions against known outcomes to determine its accuracy and reliability.

The function of the model is represented as $f(x;W,b)$. In this notation, x is the input data, and $\hat{y}$ is the output or prediction as

$$\hat{y} = f(x;W,b), \tag{2.21}$$

where W and b denote the model's weights and biases, respectively. The central objective of the training phase is to determine the optimal values for these parameters that would minimise the loss function. This optimisation can be mathematically represented as

$$(W^*,b^*) = \underset{W,b}{\operatorname{argmin}}\, L(y,\hat{y}), \tag{2.22}$$

where W^*, b^* represent the optimised weights and biasses, respectively.

In deep learning, the backpropagation algorithm computes the gradient of the loss for each model parameter. The gradients, represented as $\frac{\partial L}{\partial W}$ for weights W and $\frac{\partial L}{\partial b}$ for biases b, indicate how changes in these parameters affect the loss.

An optimiser uses these gradients to update the weights and biases during training to minimise the loss. The specific update rule of the optimiser is given by Φ and can vary based on the chosen

optimisation technique. The update rule can be expressed as

$$(W_{\text{new}}, b_{\text{new}}) = \Phi(W_{\text{old}}, b_{\text{old}}, \frac{\partial L}{\partial W}, \frac{\partial L}{\partial b}) \tag{2.23}$$

where the model convergence to a state with a minimised loss is ensured. The following section delves deeper into the various types of loss functions and optimisation strategies.

2.3.2 Loss Functions and Optimisers in Deep Learning

Deep learning models utilise loss functions and optimisers as critical components during their training. A loss function, often referred to as a cost function, is a mathematical formulation that quantifies the disparity between the model's predictions and the true outputs. As the training progresses, the primary objective is to minimise this discrepancy, ensuring that the model's predictions align closely with the actual values. This is achieved through the backpropagation process, where the gradient of the loss function with respect to each model parameter is computed. Optimisers then utilise these gradients to adjust the model's parameters in a direction that reduces the loss value. These tools not only form the backbone of deep learning techniques but also play a pivotal role in the research presented in this thesis.

The primary loss functions employed in deep learning, and also adopted in this thesis, are detailed subsequently. In the following loss functions, $y[n]$ represents the true output and $\hat{y}[n]$ denotes the model's predicted output.

- **Mean Squared Error (MSE) Loss** is prevalently used for regression tasks. The rationale behind this is that MSE provides a quantifiable measure of how close the model's continuous-valued predictions are to the actual values. By squaring the differences, it penalises larger deviations more severely than smaller ones, ensuring that the model is sensitive to larger errors. The MSE can also be interpreted from a probabilistic standpoint: if the model residuals are assumed to be Gaussian distributed with mean zero and variance σ^2, then minimizing MSE corresponds to maximizing the log-likelihood of observing the data $\{y[n]\}_{n=1}^{N}$ under the model's predictions $\{\hat{y}[n]\}_{n=1}^{N}$. Hence, the MSE is essentially the negative of the log-likelihood under Gaussian assumptions, scaled by $\frac{1}{N}$. Mathematically, for N data points, MSE is defined as

$$L_{MSE}(y, \hat{y}) = \frac{1}{N} \sum_{n=1}^{N} (y[n] - \hat{y}[n])^2. \tag{2.24}$$

- **Binary Cross-Entropy (BCE) Loss** is predominantly employed for binary classification tasks. It quantifies the difference between the true labels and the predicted probabilities. BCE is especially useful because it imposes a larger penalty when confident predictions are incorrect, making the model more cautious about its predictions. From a probabilistic perspective, if the true label $y[n]$ for each data point n is modeled as a Bernoulli random variable with success probability $\hat{y}[n]$ (the model's predicted probability), then minimizing BCE is equivalent to maximizing the log-likelihood of observing the given set of labels under the model predictions. This is because

the loss for each point involves the log-probability of the observed label under the predicted Bernoulli distribution, thus directly linking it to likelihood maximization. Mathematically, for N data points, BCE is given by

$$L_{BCE}(y,\hat{y}) = -\frac{1}{N}\sum_{n=1}^{N}\left(y[n]\log(\hat{y}[n]) + (1-y[n])\log(1-\hat{y}[n])\right). \tag{2.25}$$

- **Cross-Entropy (CE) Loss** extends the concept of BCE to multi-class classification problems. It measures the dissimilarity between the true label distribution and the predicted probabilities, penalizing predictions that are confident but incorrect. From a probabilistic standpoint, if the true labels for each data point n are modeled as a categorical distribution where only one class label is true at a time (i.e., the label vector $y[n]$ is a one-hot encoded vector representing the true class), and the predicted probabilities $\hat{y}[n]$ represent the model's estimated probabilities of each class being the true class, then minimizing CE is equivalent to maximizing the log-likelihood of observing the given set of true class labels under the model predictions. Each term of the CE loss involves the log-probability of the true class according to the predicted categorical distribution, thus enhancing the model's focus on correctly predicting the log-probability for the actual class. Mathematically, for N data points with C classes, CE is defined as:

$$L_{CE}(y,\hat{y}) = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{C} y_c[n]\log(\hat{y}_c[n]). \tag{2.26}$$

- **Soft-DTW (sDTW) Loss** refines the conventional Dynamic Time Warping (DTW) approach, tailoring it for sequence analysis. Its unique differentiable characteristic facilitates seamless integration with the backpropagation method in deep learning.

  Consider two sequences: the true sequence $y[n]$ spans from $n = 1$ to N, and the predicted sequence $\hat{y}[m]$ spans from $m = 1$ to M. The lengths of $y[n]$ and $\hat{y}[m]$ are represented by N and M, respectively.

  In traditional DTW, a matrix D is created where each element $D_{n,m}$ measures the Euclidean distance between $y[n]$ and $\hat{y}[m]$. Diverging from this, Soft-DTW introduces a matrix R. Each component of R is derived from the equation

$$R_{n,m} = D_{n,m} + \text{softmin}(R_{n-1,m}, R_{n,m-1}, R_{n-1,m-1}), \tag{2.27}$$

  where

$$\text{softmin}(a,b,c) = -\gamma\log\left(e^{-a/\gamma} + e^{-b/\gamma} + e^{-c/\gamma}\right), \tag{2.28}$$

  and γ represents a smoothing parameter. As γ approaches zero, the softmin function converges to the traditional minimum function. In essence, γ regulates the trade-off between the accuracy of the alignment and the smoothness (or differentiability) of the softmin operation.

The Soft-DTW distance between the sequences y and $\hat{y}$ is encapsulated in $R_{N,M}$. Consequently, the loss function using Soft-DTW is:

$$L_{sDTW} = R_{N,M} \tag{2.29}$$

Thanks to its design, Soft-DTW adeptly detects even minor misalignments between actual and forecasted sequences. This precision ensures minute variations in expansive sequences are recognized, enhancing the prediction accuracy of deep learning models.

Once the loss is computed, the next step involves adjusting the model's parameters to reduce this loss. This task is achieved using optimisers.

- **Stochastic Gradient Descent (SGD)** is a widely used optimisation algorithm in deep learning. It operates by iteratively updating the model parameters for each data point in the training set. This approach leads to faster convergence and efficient avoidance of local minima.

  For the sake of simplicity in the representation of parameter updates, the weights W and biases b are collectively denoted as θ. In each iteration, the model parameters θ are adjusted using the gradient of the loss function with respect to the parameters. The learning rate η determines the step size for these parameter updates. Mathematically, the update equation for SGD is given as

  $$\theta_{i+1} = \Phi(\theta_i, L(y, \hat{y})) = \theta_i - \eta \frac{\partial L(y, \hat{y})}{\partial \theta_i}, \tag{2.30}$$

  where θ_i represents the model parameters at iteration i, η is the learning rate, and $\frac{\partial L(y,\hat{y})}{\partial \theta_i}$ is the gradient of the loss function with respect to the parameters θ_i.

- **ADAM (Adaptive Moment Estimation)** is another key optimisation algorithm in deep learning. Unlike SGD that updates model parameters using only the first moment (mean) of the gradients, ADAM employs both the first moment (mean) and the second moment (uncentered variance) of the gradients to adjust the parameters. This dual-moment consideration often results in faster convergence and stability during the training process.

  As before, for simplicity in representation, the weights W and biases b are collectively denoted as θ. The ADAM optimizer updates parameters using a combination of gradient information and running averages of previous gradients. These updates are guided by hyperparameters such as

the learning rate η, and decay rates β_1 and β_2. Mathematically, the update rules for ADAM are

$$m_{t+1} = \beta_1 m_t + (1-\beta_1)\frac{\partial L(y,\hat{y})}{\partial \theta_t} \tag{2.31}$$

$$v_{t+1} = \beta_2 v_t + (1-\beta_2)\left(\frac{\partial L(y,\hat{y})}{\partial \theta_t}\right)^2 \tag{2.32}$$

$$\hat{m}_{t+1} = \frac{m_{t+1}}{1-\beta_1^{t+1}} \tag{2.33}$$

$$\hat{v}_{t+1} = \frac{v_{t+1}}{1-\beta_2^{t+1}} \tag{2.34}$$

$$\theta_{t+1} = \Phi(\theta_t, L(y,\hat{y})) = \theta_t - \frac{\eta}{\sqrt{\hat{v}_{t+1}}+\varepsilon}\hat{m}_{t+1} \tag{2.35}$$

In this context, m_t and v_t denote the first and second moment estimates of the gradients, respectively. The parameters β_1 and β_2 are exponential decay rates for these moment estimates. The term ε is a small constant added to prevent division by zero. Just as with SGD, θ_t represents the model parameters at iteration t, and η indicates the learning rate.

In deep learning, the choice of loss functions and optimisers is pivotal. They critically influence the training dynamics and the final model performance. While a plethora of options exists, the Mean Squared Error, Binary Cross-Entropy, and Soft-DTW losses, in conjunction with the Schotastic Gradient Descent and ADAM optimisers, are central to the methodologies deployed in this research.

2.3.3 Neural Networks Layers

This section presents the foundational components of the deep learning model function as shown in (2.21). Neural networks comprise multiple layers and building blocks, each tailored to specific data types and objectives. The primary layers employed in this research are briefly described as follows.

- **Fully Connected (FC) Layer** serves as a primary component in neural networks where every neuron connects to all neurons from the preceding layer. It computes outputs by applying a weighted linear combination of its inputs, combined with biases, and subsequently processes this through an activation function. For a layer with N input neurons and M output neurons, the mathematical representation is:

$$y = \sigma(Wx+b), \tag{2.36}$$

where W is an $M \times N$ weight matrix, b is an $M \times 1$ bias vector, and σ denotes the activation function applied element-wise.

- **Sigmoid Activation Function** scales its input values to lie between 0 and 1. This mapping is achieved using the mathematical relationship

$$y = \frac{1}{1+e^{-x}}, \tag{2.37}$$

where y is the output after applying the sigmoid function to the input x.

- **ReLU Activation Function** applies an element-wise operation that outputs the input value if positive; otherwise, zero, which is represented as

$$y[i] = \max(0, x[i]) \quad \text{for } i = 0, 1, \ldots, |x| - 1, \tag{2.38}$$

where $y[i]$ is the output for the i^{th} element after applying the ReLU function to the i^{th} element of input x and $|x|$ denotes the length of vector x.

- **Softmax Activation Function** computes the exponentials for each input element and then normalises these values, producing a probability distribution over multiple classes. The mathematical representation of this process is

$$y[i] = \frac{e^{x[i]}}{\sum_{j=0}^{|x|-1} e^{x[j]}} \quad \text{for } i = 0, 1, \ldots, |x| - 1, \tag{2.39}$$

where $y[i]$ denotes the probability associated with the i^{th} class. The denominator guarantees the summation of probabilities over all classes is one.

- **Tanh Activation Function** compresses the input values into a range between -1 and 1, providing a zero-centered output, which often aids optimization. The mathematical expression for the tanh function is

$$y[i] = \frac{e^{x[i]} - e^{-x[i]}}{e^{x[i]} + e^{-x[i]}} \quad \text{for } i = 0, 1, \ldots, |x| - 1, \tag{2.40}$$

where $y[i]$ represents the transformed value corresponding to the input $x[i]$. Due to its zero-centered nature, the tanh activation function is preferred over the sigmoid function in certain scenarios, especially for intermediate layers in deep networks.

- **1D Convolutional (Conv) Layer** performs a one-dimensional (1D) convolution operation by sliding a kernel, also known as a filter, over the input data. Given an input sequence x of length I and a kernel k of size K, the convolutional operation generates outputs by iteratively applying the kernel with a defined stride S. To ensure that the input and output sequence sizes remain consistent throughout this work, appropriate zero padding is applied. Therefore, even though the inherent output sequence length O would be $O = \frac{I-K}{S} + 1$, zero padding is added to maintain $O = I$.

  The layer employs F distinct kernels. Each kernel is designed to capture different features or patterns from the input sequence. By having multiple kernels, the network can identify a variety of features, ensuring versatility. For instance, while one kernel might recognise a specific pattern, another could detect a different sequence feature. Such a design is pivotal, especially in tasks like sequence analysis where multiple patterns could emerge.

  For each applied kernel, a unique output sequence is generated. Utilizing F kernels results in the production of F distinct output sequences. These sequences collectively form a 2D output tensor with dimensions $I \times F$. The convolution operation, when applied to a specific kernel f at

a particular location i in the output tensor, is represented by

$$y[f,i] = \sum_{j=0}^{K-1} k[f,j] \cdot x[i \times S + j], \tag{2.41}$$

where S is the stride or step size taken when shifting the kernel across the input. The kernel index, f, ranges from 1 to F.

In situations where the input tensor x encompasses more than one channel, its dimensions become $I \times C$, with I indicating the sequence length and C representing the number of channels. Concurrently, the kernel tensor, signifying the weights applied during the convolution process, adopts a shape of $F \times C \times K$, where F is the count of filters. In this multichannel context, the convolution operation modifies to

$$y[f,i] = \sum_{c=0}^{C-1} \sum_{j=0}^{K-1} k[f,c,j] \cdot x[i \times S + j, c]. \tag{2.42}$$

For the remainder of this thesis, a 1D convolution operation is denoted as $F@Conv(KxS)$, where F symbolises the number of distinct kernels, K indicates the kernel size, and S signifies the stride.

- **1D Separable Convolution (SConv) Layer** performs a two-stage convolution operation on the input data. The motivation behind this layer is to achieve computational efficiency, especially when the input has a significant number of channels.

  The first stage involves a depthwise convolution. Contrary to conventional convolutional processes that employ the same kernel across all input channels, depthwise convolution applies a distinct filter to each input channel. Given an input tensor $x[i,c]$ with a sequence length of I and chanel size of C, and considering a depthwise kernel $k_{\text{dw}}[j,c]$ of dimensions (K,C), the output from the depthwise convolution for each channel, represented as

$$y_{\text{dw}}[i,c] = \sum_{j=0}^{K-1} k_{\text{dw}}[j,c] \cdot x[i \times S + j, c], \tag{2.43}$$

  where i is the index in the output tensor, j is the index of the kernel, c designates the channel index ranging from 1 to C, and S represents the stride length.

  In the second stage, the pointwise convolution combines the channels outputted by the depthwise convolution. This is achieved by using F distinct $1 \times C$ kernels. For each filter f, ranging from 1 to F, the pointwise convolutional output $y[i,f]$, also the SConv layer output is calculated as

$$y[i,f] = \sum_{c=1}^{C} k_{\text{pw}}[c,f] \cdot y_{\text{dw}}[i,c], \tag{2.44}$$

where $k_{\text{pw}}[c, f]$ denotes the weight of the pointwise kernel associated with channel c for the filter f.

Decomposing the convolution in this manner allows the SConv layer to drastically diminish the number of parameters and computational overhead, particularly when the input has a high channel count C. This design offers an efficient counterpart to traditional convolution operations, enhancing the model's efficiency.

For subsequent references in this thesis, a 1D SConv operation will be denoted as $F@SConv(K,P)$, where F symbolises the number of unique kernels, K indicates the kernel size, and P designates the zero padding size. The stride value is always kept as one for 1D Sconv operations in this thesis.

- **Batch Normalization (BN) Layer** normalises the activations of a given layer across each batch. Specifically, it transforms the activations to ensure a near-zero mean and a standard deviation close to 1. This stabilising operation has become pivotal in deep learning as it helps in accelerating convergence and alleviating issues related to internal covariate shift.

  Given activations x, the normalized output y is computed using:

$$y = \frac{x - \mu}{\sqrt{\sigma^2 + \varepsilon}} \times \gamma + \beta, \tag{2.45}$$

  where μ denotes the mean of x, σ^2 stands for its variance, γ is a scale factor, β is an offset factor, and ε is a small constant added for numerical stability.

  The scale and shift parameters, γ and β respectively, are learned during the training process, allowing the layer to adaptively scale and shift the activations as needed. Such adaptability ensures that the layer doesn't overly constrain the capacity of the network.

  This layer is often utilised after 1D Conv or FC layers and before non-linear activation functions to ensure a smooth training process and stable convergence.

- **Adaptive Batch Normalization (AdaBN) Layer** refines the standard BN method to better suit domain adaptation tasks. While BN typically uses the current training mini-batch to compute mean and variance statistics, this might not be optimal when transitioning from a source to a target domain, especially in unsupervised domain adaptation scenarios.

  AdaBN bridges this gap by adjusting these statistics during inference. Rather than sticking to statistics from the source domain (training data), it utilises the statistics from the target domain (testing data). In essence, the mean (μ) and variance (σ^2) determined from the source are replaced by those from the target domain when adapting.

  This proactive recalibration by AdaBN ensures that the model remains consistent across both domains, boosting its ability to generalise and effectively transfer knowledge between them.

- **1D Max Pooling Layer** employs a down-sampling strategy by sliding a window, sized K, over the input using a stride S. At each position, it selects the highest value within this window to

form the output. For an input tensor x, the output at location i is computed as

$$y[i] = \max_{j=0}^{K-1} x[i \times S + j], \tag{2.46}$$

where both K and S represent the window size and stride length, respectively.

- **1D Average Pooling Layer** similarly down-samples the input using a window of size K and stride S. However, instead of capturing the maximum value, it calculates the average of all values within the window. Given an input tensor x, the output for a specific location i is

$$y[i] = \frac{1}{K} \sum_{j=0}^{K-1} x[i \times S + j], \tag{2.47}$$

where K is the window size and S signifies the stride length.

- **1D Adaptive Average Pooling Layer** automates the process of down-sampling the input tensor to generate an output of a predetermined target size. Contrary to the conventional average pooling, where the kernel size and stride are explicitly specified, this layer dynamically computes them based on the input size and the desired output size.

  For an input tensor x of length I and a designated output size O, the kernel size K and stride S for the pooling are derived as:

$$K = \left\lceil \frac{I}{O} \right\rceil, \tag{2.48}$$

$$S = \left\lfloor \frac{I}{O} \right\rfloor. \tag{2.49}$$

  Using these determined values of K and S, the adaptive average pooling for a particular location i in the output tensor is:

$$y[i] = \frac{1}{K} \sum_{j=0}^{K-1} x[i \times S + j]. \tag{2.50}$$

  The standout advantage of adaptive pooling lies in its adaptability. It equips the neural network with the capacity to accommodate tensors of varied sizes, ensuring a uniform output size without manual tweaks to the kernel or stride parameters.

- **Recurrent Neural Network (RNN) Layer** processes sequential data by maintaining a dynamic state that encapsulates the historical context of the sequence. Given an input sequence $x = (x_1, x_2, \ldots, x_T)$, where each $x[n]$ symbolizes the vector at the time step n and considering an initial state $h[0]$, the RNN iteratively determines the states and outputs at each time step.

For a particular time step n, the state $h[n]$ and the resultant output $y[n]$ are derived as:

$$h[n] = \sigma\left(W_{hh}h[n-1] + W_{xh}x[n] + b_h\right) \tag{2.51}$$

$$y[n] = W_{hy}h[n] + b_y \tag{2.52}$$

Here, σ signifies the activation function (commonly tanh or sigmoid), while W_{hh}, W_{xh}, b_h, and b_y represent the associated weight matrices and bias vectors.

Though the RNN layer inherently captures past information, its memory span can be constrained, often due to complications like the vanishing gradient problem. To navigate these issues, advanced variants such as LSTM and GRU are typically employed.

- **Long Short-Term Memory (LSTM) Layer** addresses the limitations of standard RNNs by introducing memory cells and three gating mechanisms: input, forget, and output gates. These structures enable the LSTM to learn long-term dependencies and combat the vanishing and exploding gradient problems inherent to simple RNNs.

  Given an input sequence $x = (x_1, x_2, \ldots, x_T)$, where each $x[n]$ is the vector at time step n, and with initial states $h[0]$ and $c[0]$ (the cell state), the LSTM computes the states and outputs for each time step iteratively.

  For each time step n, the various components are calculated as:

$$f[n] = \sigma\left(W_f \cdot (h[n-1], x[n]) + b_f\right) \tag{2.53}$$

$$i[n] = \sigma\left(W_i \cdot (h[n-1], x[n]) + b_i\right) \tag{2.54}$$

$$\tilde{c}[n] = \tanh\left(W_c \cdot (h[n-1], x[n]) + b_c\right) \tag{2.55}$$

$$c[n] = f[n] \circ c[n-1] + i[n] \circ \tilde{c}[n] \tag{2.56}$$

$$o[n] = \sigma\left(W_o \cdot (h[n-1], x[n]) + b_o\right) \tag{2.57}$$

$$h[n] = o[n] \circ \tanh(c[n]) \tag{2.58}$$

  Here, σ is the sigmoid activation function, $\circ$ denotes element-wise multiplication, and the weight matrices W_f, W_i, W_c, W_o and bias vectors b_f, b_i, b_c, b_o are the trainable parameters associated with the forget, input, cell, and output gates, respectively.

  The hidden state $h[n]$ at each time step encapsulates the abstract representation of the input sequence up to that time step. This state can be transformed into an output $y[n]$ using the equation

$$y[n] = W_y \cdot h[n] + b_y, \tag{2.59}$$

  where W_y represents the weight matrix and b_y is the associated bias vector. In deep learning architectures, the hidden state $h[n]$ serves as a foundational component, often being used as input for subsequent layers, be they recurrent, convolutional, or fully connected. On the other hand, the output $y[n]$ is typically employed for direct tasks such as classification or regression. For models incorporating multiple recurrent layers, the hidden state from one LSTM layer

can seamlessly transition to serve as input for the next LSTM layer. Furthermore, in specific applications where only the final output holds relevance, the last hidden state $h[T]$ offers a summarised representation of the entire sequence.

- **Bidirectional Long Short-Term Memory (BLSTM) Layer** processes an input sequence in both forward and backward directions. By doing so, it captures both past and future context for each time step. This results in two sets of hidden states, $\overrightarrow{h[n]}$ and $\overleftarrow{h[n]}$, which can be concatenated or summed to form the final output at each time step. It's particularly beneficial for tasks where future context provides essential information, such as in sequence labeling and pattern extraction.

2.3.4 Introduction to Performance Metrics

After the training phase, the model undergoes testing using an unseen dataset referred to as the test dataset. Performance metrics serve as tools to evaluate the effectiveness and reliability of the trained model, facilitating a comparison with the state of the art. These metrics offer quantifiable measures to investigate the accuracy of a model's predictions in contrast to the actual outcomes. Proper assessment with these metrics highlights areas of improvement, assists in model selection, and guides the decision-making process. This thesis utilizes the performance metrics detailed in this section.

The confusion matrix is a vital evaluation tool for classification models, comparing actual classes with predicted outcomes in a structured table.

	Predicted: Positive	**Predicted: Negative**
Actual: Positive	True Positives (TP)	False Negatives (FN)
Actual: Negative	False Positives (FP)	True Negatives (TN)

In this matrix, True Positives (TP) are the cases where the model correctly predicts a positive outcome. False Negatives (FN) occur when the model inaccurately labels a positive case as negative. False Positives (FP) represent instances where a negative case is wrongly predicted as positive. True Negatives (TN) are when the model accurately predicts a negative outcome. The confusion matrix is invaluable in offering a clear overview of a model's classification accuracy, allowing for a comprehensive analysis of its performance.

Using the confusion matrix, we can derive several key performance metrics that assess the effectiveness of a model. Accuracy gauges the model's overall correctness, and it is defined as

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{2.60}$$

Precision, also called the positive predictive value, assesses the model's exactness by measuring the correctness of positive predictions such that

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \tag{2.61}$$

Recall, also referred to as Sensitivity, evaluates the model's ability to identify all relevant instances, expressed as

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{2.62}$$

False Omission Rate (FOR) is a measure that informs us of the proportion of negative instances that were incorrectly identified as positive, given by

$$\text{FOR} = \frac{\text{FN}}{\text{TN} + \text{FN}}. \tag{2.63}$$

False Discovery Rate (FDR) determines the proportion of positive identifications that were incorrect, expressed as

$$\text{FDR} = \frac{\text{FP}}{\text{TP} + \text{FP}}. \tag{2.64}$$

Negative Predictive Value (NPV) assesses the correctness of negative predictions, defined as

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}}. \tag{2.65}$$

Specificity, also known as True Negative Rate, gauges the model's ability to rightly identify negative instances, expressed as

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \tag{2.66}$$

The F1 Score serves as a balanced metric between precision and recall, particularly crucial in scenarios with uneven class distributions. It can be expressed as

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{2.67}$$

Two specialised adaptations of the F1 Score, Macro and Micro, address the nuances in multi-class settings. The Macro F1 Score is given by

$$\text{Macro F1 Score} = \frac{1}{N} \sum_{i=1}^{N} \text{F1 Score}_i, \tag{2.68}$$

where i indexes each of the N distinct classes in the dataset. This version computes the F1 Score for each class individually and then averages these values. It is particularly useful when performance needs to be understood across varying classes.

On the other hand, the Micro F1 Score is represented as

$$\text{Micro F1 Score} = \frac{\sum_{i=1}^{N} \text{TP}_i}{\sum_{i=1}^{N} \text{TP}_i + 0.5(\sum_{i=1}^{N} \text{FP}_i + \text{FN}_i)}, \tag{2.69}$$

which aggregates the outcomes from all classes to determine a global metric. By providing equal weight to each instance, the Micro F1 Score becomes particularly pertinent in multi-label scenarios.

Chapter 3

Datasets Utilised: Acquisition, Description, and Experimental Context

In this study, experiments are conducted using two distinct datasets: Ninapro, denoted as D_n, and Football, denoted as D_f. Each dataset comprises pairs (x^i, y^i), where x^i represents the recorded sEMG signal, y^i indicates the true label, and i denotes the index of the data sample. The details of the Ninapro dataset, D_n, are outlined in Section 3.1. Similarly, the specifics of the Football dataset, D_f, are detailed in Section 3.2. A thorough comparison between the two datasets, D_n and D_f, is presented in Section 3.3. This comparison encompasses the similarities and differences between the datasets, clarifying their collaborative suitability for the methodologies and analyses elaborated in the subsequent chapters of this research.

3.1 Ninapro Dataset (D_n)

The Ninapro dataset, denoted as D_n, serves as a trusted benchmark in the evaluation of motion classification models. Available online, this study utilises the DB2 subset of Ninapro, which includes data for 49 hand gestures collected from 40 healthy participants [9]. Each gesture in D_n comprises six repetitions, each lasting five seconds, separated by a three-second rest interval. The data is produced by 12 sEMG sensors. Specifically, eight sensors are evenly distributed around the forearm at the level of the radio humeral joint. Additionally, two sensors are assigned to the flexor and extensor digitorum superficialis, while the remaining two are attached to the biceps brachii and triceps brachii. All sensors are consistently positioned on the right hand for every participant. The sEMG recordings employ Delsys Trigno sEMG sensors, characterised by bipolar differential electrodes, a 2 kHz sampling rate, 16-bit ADC resolution, and a 20-450 Hz bandpass filter [9]. The comprehensive specifications of the Delsys Trigno sEMG sensors is provided in Appendix A.

The gestures within the DB2 Ninapro dataset are further categorised into three subsets: B, C, and D. Specifically, Subset B, $D_{n,b}$, encompasses 8 isometric and isotonic hand configurations, alongside 9 basic wrist movements; Subset C, $D_{n,c}$, details 23 grasping and functional movements; and Subset D, $D_{n,d}$, describes 9 finger force patterns.

For the purposes of this thesis, both the complete Ninapro dataset D_n and its subsets $D_{n,b}$ and $D_{n,d}$ are employed. In alignment with convention, repetitions 1, 3, 4, and 6 are designated for training, while repetitions 2 and 5 are reserved for testing.

3.2 Football Dataset (D_f)

Our Football dataset, D_f, represents sEMG data collected from 45 athletes across two English Premier League (EPL) clubs spanning two and a half competitive seasons. Unlike many datasets, D_f captures data during the athletes' real-life exercises, making it unique in its representation of less controlled environments. On the day following each match (MD+1), players complete 9 exercises in a random order. These exercises have varied resting periods, sometimes incorporating walking and stretching.

The sensors, positioned bilaterally, measure 8 lower extremity muscle groups: biceps femoris (long head), semitendinosus, adductor longus, and soleus. The employed Neurocess sEMG sensors come with bipolar differential electrodes and include features such as a 1 kHz sampling rate, a 20-450 Hz bandpass filter, and a 16-bit ADC resolution. Besides sEMG data, the dataset also contains the HASI information for every athlete. For detailed specifications of the Neurocess sEMG sensors utilised in this study, refer to Appendix A.

The lead physiotherapists of the respective clubs supervised the experimental protocols and sensor placements. The physiotherapists who conducted the experiments were trained on how to position the sensors to the muscle groups and the sensor alignments to maintain consistency over various data collections, highlighting the emphasis on quality control. sEMG recordings took place at the training grounds, having secured the ethical clearance from King's College London (reference MRSP-20/21-20996). All personal data undergoes pseudonymisation in line with privacy standards. The link between identifiable athlete information and their pseudonyms remains securely stored by the clubs. This pseudonymisation approach is crucial for the dataset since clubs utilise the data and record injuries per player. Keeping injury data while maintaining athlete confidentiality necessitates pseudonymisation. Subsequent subsections provide in-depth discussions on sensor placement and the drills recorded.

3.2.1 Optimal sEMG Sensor Placement

The sEMG sensor placement is one of the key factors that influence the sEMG signal [113, 59]. There are three significant things to raise attention in sensor placement; these key factors are orientation of the electrodes, position of the sensor on the muscle, and knowledge of individual innervention zone (IZ) [113].

As recommended by Surface EMG for Non-Invasive Assessment of Muscles (SENIAM) [63], the electrode orientation should be parallel with the muscle fibres. As indicated by [113] the proper allignment can only be obtained by the orientation in which a symmetric waveform pattern of the action potential propagating in the two directions between the end plate and the two tendon regions

is observed. For the reproducibility of the experiments, specific sensor positions for each muscle are detailed as follows.

- **Biceps Femoris Long Head:** Sensors are positioned at the 50% mark along the line connecting the ischial tuberosity to the lateral epicondyle of the tibia.
- **Semitendinosus:** Sensors are positioned at the midpoint, or 50%, on the line extending from the ischial tuberosity to the medial epicondyle of the tibia to monitor the semitendinosus muscle.
- **Adductor Longus:** Sensors are placed at the 66% point on the line running from the pubic tubercle to the medial epicondyle of the femur, directly over the muscle belly.
- **Soleus:** Sensors are positioned at the 66% mark on the line linking the medial condyle of the femur to the medial malleolus.

Bipolar differential electrodes, positioned equidistantly on either side of the IZ, may degrade the sEMG signal quality [113]. Such placement can cause the electrodes to cancel out the action potentials (APs) as symmetric waveforms start from the IZ and progress towards the tendons [102]. When electrodes are directly over the IZ, the action potentials might be attenuated, resulting in a weaker sEMG signal or a noise-like output [113].

SENIAM provides general recommendations for sensor placement, but individual variations may require adjustments to these guidelines. The IZ location can vary across individuals, necessitating a check to ensure the SENIAM-advised position matches an athlete's specific IZ. Moreover, the muscle structure of elite athletes might be influenced by their history of muscle injuries or surgeries. These structural changes underline the importance of personalised sensor placement, which might diverge from the standard SENIAM recommendations.

To determine whether an athlete needs a modification from the standardised SENIAM recommendation locations, a 5 sEMG electrode array is used, aligned with the direction of each muscle fibre considered in this study. The array, consisting of 5 sensors with a distance between electrode pairs of 8mm, is situated around the SENIAM recommended position as previously detailed. Given that each sensor spans 33mm, a total area of 165mm is monitored using this electrode array. Athletes perform maximal isometric contractions for each muscle. If the sEMG signal from the SENIAM-recommended position displays either a noise-like signal or a significant drop in amplitude, it suggests that the IZ is aligned with SENIAM's guideline. In such cases, the optimal sensor placement is chosen based on the location delivering the highest sEMG amplitude.

3.2.2 Description of Selected Drills

Athletes are directed to execute a series of hamstring, adductor, and soleus exercises to evaluate muscle activation patterns through sEMG recordings. The details of the selected drills are listed as follows.

1. 60° **Adductor Squeeze:** Athletes lie supine (on their back) with their hips flexed at 60°. They then squeeze a foam roll placed between their knees, ensuring consistent pressure throughout the

exercise. Athletes gradually increase the intensity of the squeeze to reach a maximal contraction over two seconds. Once at maximum, they hold this intense squeeze isometrically for three seconds, before gradually relaxing over the subsequent two seconds. This exercise primarily targets the adductors, particularly the adductor longus, and also engages the hamstrings due to the stabilizing requirements of the hip joint.

Fig. 3.1 Illustration of an athlete performing a hamstring claw exercise. sEMG sensors, attached bilaterally on four muscle groups using adhesive medical tapes, are shown. In this figure, only the sensors on the left soleus and hamstring muscles are visible. Sensors on the biceps femoris long head and semitendinosus are marked with a dashed black line, while the sensor on the soleus is indicated by a straight black line.

2. **Hamstring Claw:** As illustrated in Figure 3.1, athletes lie prone (face down) over a supporting box, setting their hips at 45° flexion and knees at 15° flexion. They extend the thigh by lifting the leg, ensuring a stable and consistent posture throughout. This is a single-leg exercise, performed individually for each leg, and the knee angles are always kept the same. The exercises are realized within a Kanga-tech force frame, where the contraction is built up against a fixed resistance. The resistance force frame pad is positioned just above the Achilles tendon. The contraction is built up gradually over two seconds to maximum intensity and held isometrically

at this peak for two seconds. The exercise is performed for a total of three repetitions. The primary muscles targeted are the hamstrings, particularly the biceps femoris, with the adductors providing additional support for stabilization.

3. **Hip Extension:** Athletes position themselves prone, with the pelvis resting on a mat and their hips at 30° flexion and knees at 60° flexion. This setup ensures that the hamstrings are adequately pre-stretched to maximize engagement. This is a single-leg exercise, performed individually for each leg, and the knee angles are always kept the same. The exercises are realized within a Kanga-tech force frame, where the contraction is built up against a fixed resistance. The resistance force frame pad is positioned just above the above popliteal crease. Then, they extend the hip to elevate the thigh, ensuring a controlled and smooth movement throughout.The contraction is gradually built up over two seconds, held at maximal intensity isometrically for three seconds, and then relaxed over two seconds. Athletes repeat this sequence for three repetitions. The hamstrings, including the semitendinosus and biceps femoris, are the primary muscles engaged, followed by the adductors for hip stabilization. This drill is particularly useful for targeting the posterior chain.

4. **Hip Flexion:** Athletes lie supine, with hips flexed at 60° and feet flat on the ground and knees flexed at 90°. This is a single-leg exercise, performed individually for each leg, and the knee angles are always kept the same. The exercises are realized within a Kanga-tech force frame, where the contraction is built up against a fixed resistance. The resistance force frame pad is positioned just above the knee cap. Then, they lift the heel of the test leg, maintaining the hip and knee flexion angle. The contraction is built up over two seconds, held isometrically at maximum intensity for three seconds, and then relaxed over another two seconds. This exercise is repeated for three repetitions. It primarily targets the adductors, which are highly recruited as stabilizers, followed by the hamstrings.

5. **Sitting Soleus Raise:** Seated on a 45cm box with their knee flexed at 90°, athletes exert maximal force to elevate the heel, ensuring the ball of the foot remains grounded. The contraction intensity is built up gradually over two seconds, reaching its peak by this point. Athletes then hold this maximum contraction isometrically for three seconds, before easing the tension and relaxing over the next two seconds. This pattern is repeated for three repetitions in total. The primary muscle targeted by this exercise is the soleus, with the hamstrings also being engaged.

6. **Single Leg (SL) Bent-Knee Soleus Raise:** Standing on a single leg, athletes bend the supporting knee to 30° and elevate the heel, ensuring stability throughout the movement. This exercise comprises only isotonic contractions and takes approximately two seconds to complete a single repetition. Athletes perform six repetitions on each leg. The primary muscle engaged during this exercise is the soleus, with the hamstrings also receiving secondary engagement to stabilize the knee joint.

7. **SL Elevated Glute Bridge:** Lying supine, with knees flexed at 90° and feet positioned on a 45cm box, athletes raise their hips off the floor to activate the hamstring muscles. Upon reaching

the peak bridge position, athletes perform a maximal isometric contraction for two seconds, then return to the starting position. One repetition takes approximately three seconds. Athletes complete three repetitions on each leg. The primary focus of this exercise is on the hamstrings followed by adductor muscles.

8. **Isometric Prone Squeeze at** 0°**:** Athletes lie prone with their knee fully extended at 0°. An examiner applies resistance at the ankle, against which the athletes contract their hamstring muscles. The contraction is held isometrically at maximal intensity for three seconds. This procedure is repeated for a total of three repetitions on each leg. The primary muscle engaged during this exercise is the hamstrings, with secondary engagement of the adductors.

9. **SL Prone Curls:** Starting in a prone position with a fully extended knee, athletes curl the knee to 90° and then revert back to the starting position. It is crucial for the athletes to ensure the hip remains stable throughout the movement. Each repetition takes approximately one second to complete, and no isometric holds or maximal contractions are applied. The exercise is performed using the athlete's own body weight, without any additional resistance. Athletes complete a total of 15 repetitions for each leg.

3.3 Comparative Analysis of D_n and D_f: Suitability for Thesis Scope

In comparing the datasets D_f and D_n, it is evident that they exhibit both similar and distinct characteristics. These aspects are crucial for understanding the influence of these datasets on the proposed methodologies in this research and for demonstrating their suitability for the study's objectives. The similarities between the two datasets are detailed as follows:

- Both D_f and D_n adhere to the SENIAM standard for data collection, ensuring controlled and consistent sensor placement procedures across both datasets.
- As illustrated in Figure 3.5 and Figure 3.4, both datasets encompass multichannel data, simultaneously collected from various muscle groups. This setup offers extensive insights into the recruitment patterns of different muscles.
- Both datasets include isometric and isotonic contractions, providing comprehensive sEMG data of various muscle contractions under distinct configurations. D_n predominantly comprises isometric contractions with minimal movement; however, it also includes exercises such as wrist supination and pronation, which involve isotonic contractions of various muscles. In isometric contractions, muscles maintain consistent tension without changing the muscle fibre length, whereas isotonic contractions involve muscle fibres either lengthening under load (eccentric) or shortening under load (concentric). Similarly, D_f encompasses both isometric contractions, such as adductor and prone squeezes at constant degrees, and isotonic exercises, like glute bridges or prone curls. Specific lower extremity activities within D_f demonstrate dynamic isotonic

contractions, alternating between eccentric and concentric phases. During these D_f exercises, there are continual changes in the angles of the knee, ankle, and hip joints. These variations are associated with fluctuations in muscle recruitment patterns and alterations in muscle fibre lengths, showcasing the complex and dynamic nature of these exercises.

The datasets D_f and D_n also exhibit significant differences, each uniquely tailored to specific contexts and subject profiles. The distinctions between them are detailed further as follows:

- The sEMG sensors used for collecting data in D_f and D_n are validated in highly respected academic journals [114, 10], yet they differ in hardware, sampling rate, and ADC resolution (see Appendix A). These differences in the sensors' technological specifications may lead to variations in the quality and type of data captured in each dataset.

- D_n encompasses data collected from a diverse sample of both male and female participants, while D_f consists of data from elite athletes, recognized internationally for their high level of proficiency in football. This distinction underscores the variability in physical capabilities and conditions of the subjects represented in each dataset.

- D_f encompasses data on bilateral muscle groups, offering a thorough perspective on muscular activity. Conversely, D_n is unilaterally focused on the muscle groups associated with the dominant hand. This divergence in the coverage of muscle groups between the two datasets indicates their distinct scopes and targeted applications.

- D_f is collected in dynamic environments characterized by athletes engaging in high-motion exercises with significant ground impact and explosive isotonic muscle contractions. This setting, aimed at enhancing athletic performance, results in the inclusion of MIA within the D_f dataset. An example of such MIA contamination is provided in Figure 3.3. In contrast, D_n is compiled in more controlled environments, focusing on exercises that involve predominantly stationary positions of hands, wrists, arms, and fingers and are characterized by isometric muscle activities along with occasional low-motion, controlled isotonic contractions. These exercises in D_n are designed not for maximal exertion but to facilitate precise prosthetic control. Such controlled conditions in D_n effectively preclude the presence of MIA.

- The dataset D_f differs from D_n in terms of the health profile of the subjects involved. Specifically, D_f includes athletes generally in good health but may have undetected injuries or muscle damage due to their active participation in competitive sports. Though not apparent during screening, such underlying conditions could lead to muscle anomalies on the verge of injuries. Consequently, the sEMG data from D_f sometimes include screenings conducted during periods when an athlete is at a heightened risk of injury, capturing data from muscles that are in a pre-injury state. This suggests that a subset of the sEMG data in D_f might reflect suboptimal health conditions, documenting the physiological state of athletes on the threshold of sustaining injuries shortly after the screening. In contrast, D_n involves subjects confirmed to be free from any health conditions or injuries, ensuring a more uniform health baseline for data collection.

- In addition to the differences in health profiles, D_f and D_n datasets are also distinct in the physical conditioning of their subjects prior to data collection. In D_f, athletes undergo intensive physical activity a day before each measurement session. This results in varying degrees of fatigue among subjects during data collection, adding a layer of variability to the sEMG data. Moreover, the exact level of fatigue in each athlete is not quantifiable within D_f, as the dataset does not include specific details on the effort exerted or the duration of physical activity. In contrast, D_n involves a screening process where subjects are not subjected to high physical load before measurement, thereby eliminating the effect of prior muscular fatigue on the sEMG data. This controlled approach in D_n ensures a more consistent physiological state among subjects at the time of data collection. Due to increased fatigue, the sEMG recordings in Figure 3.6 show that the muscles of the left adductor longus and left biceps femoris are not activated, resulting in a bilateral imbalance. This weakness is compensated by heightened activation of the left soleus and constant tension in the semitendinosus. This example underscores the high inter-session variability in D_f, as it continuously captures data under varying physical load and fatigue conditions.

- D_n is characterised by single-session data collection from various subjects, meaning each subject's data is captured just once, without any follow-up sessions. This approach results in the absence of inter-session data in the D_n dataset, thereby lacking variations that could arise from changes the subject's physiological state, such as fatigue. In contrast, D_f involves continuous data collection from the same subjects over multiple sessions. This method captures the inter-session variations, providing a more comprehensive view of each subject's physiological changes across different sessions.

- D_f is collected in realistic sports training environments, closely resembling the natural dynamics of actual sports scenarios. Within these settings, athletes partake in various activities, such as walking, jumping, or stretching, during the interludes of their training drills. This introduces an additional layer of complexity in accurately distinguishing exercise-related muscle activations from incidental ones. Moreover, the intervals between drills or repetitions in D_f are variable, aligning with the unstructured nature of real-life sports. In contrast, D_n is compiled in more structured laboratory settings, where participants' movements between different exercises or motions are systematically regulated, and break durations are uniformly enforced. As a result, the data stream from D_f exhibits greater complexity, closely reflecting real-life, uncontrolled conditions. Figure 3.2 presents the sEMG stream from the D_f dataset, alongside the accurately marked drill times, as indicated by the true labels. Conversely, D_n is marked by a higher degree of control, ensuring that irrelevant muscle contractions and the timing of breaks and drills are consistently standardized.

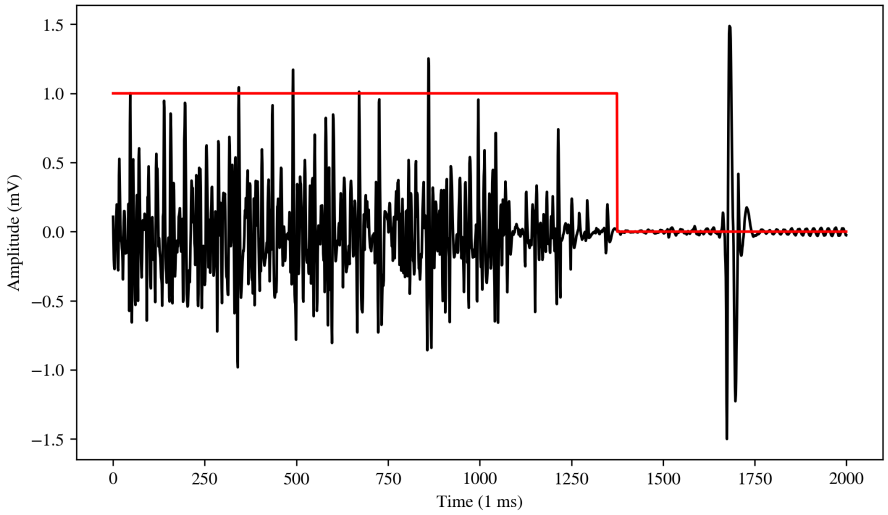
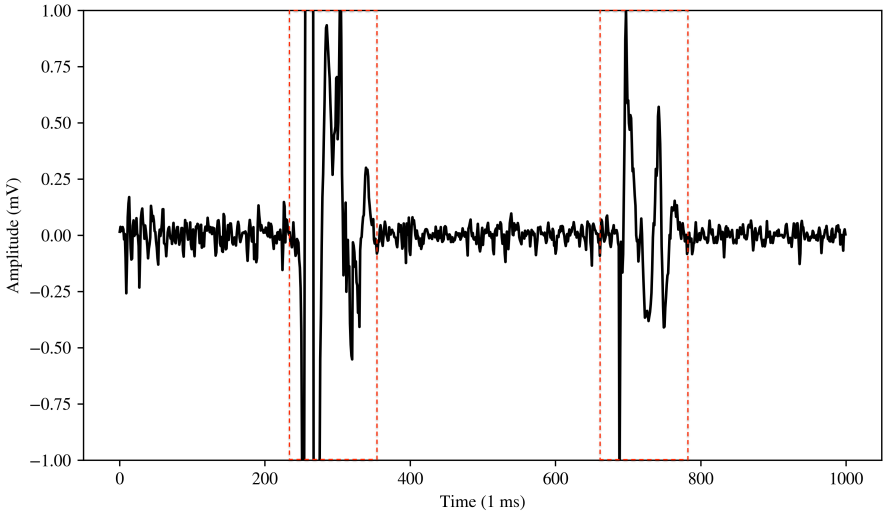


Fig. 3.2 The illustrated plot is a segment of an sEMG recording from the dataset D_f, where the sEMG stream is represented in black, displaying both contraction phases and subsequent rest periods. The x-axis represents time, while the y-axis indicates the amplitude. Overlaying the sEMG data is a label line in red, indicating the true drill times as labeled by physiotherapists using the Neurocess software. Notably, in the offset region of the plot, there is an appearance of MIA, which poses a potential challenge for onset detection. This is due to its higher peak-to-peak voltage compared to the muscle contractions, underscoring the complexities involved in accurately interpreting sEMG signals in the presence of such artefacts.

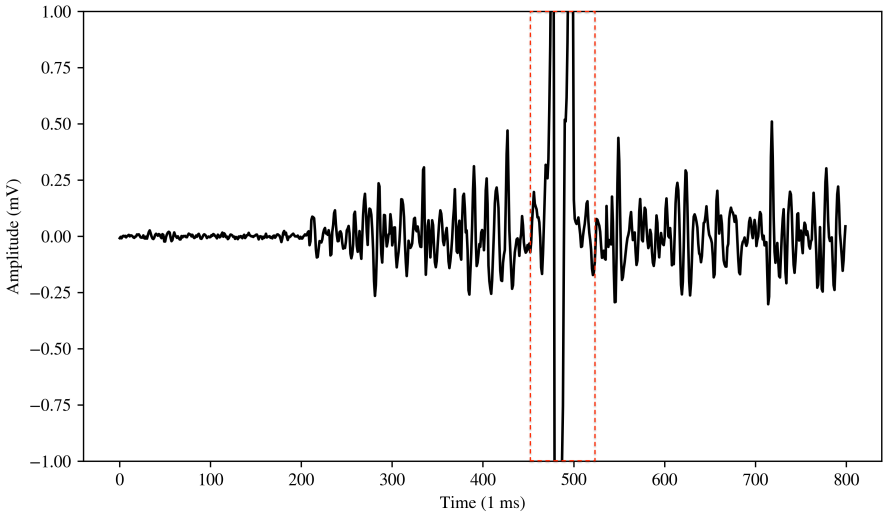
Given the comparative analysis of the datasets D_f and D_n, it is evident that they collectively offer substantial value to this research. The dataset D_f is particularly indispensable for the core objectives of this research. Collected specifically in real-life sports environments, D_f focuses on the muscles prominent to this study, particularly concerning Hamstring Adductor and Soleus Injury (HASI). It encompasses data from elite athletes, including pre-injury and injury instances alongside their healthy states. This dataset forms an ideal benchmark for testing and developing the proposed deep learning framework aimed at predicting HASI in sports through regular sEMG recordings of athletes. Its real-life, uncontrolled collection environment, as previously outlined, renders D_f vital for testing the drill time labelling model in Chapter 4 and the motion classification in Chapter 6.

The dataset D_n plays a pivotal role in this research, primarily serving as a validation set for the motion classification model outlined in Chapter 6. As a dataset that is both publicly available and extensively benchmarked, D_n offers a solid foundation for contrasting our proposed motion classification methodology with existing models in the field. Characterized by a wide range of isometric and isotonic contractions captured through multichannel sEMG sensors, D_n possesses key similarities with the D_f dataset, enhancing its applicability and value to our study. Particularly in the latter sections of Chapter 6, D_n is integral to the few-shot learning approach. Here, the fundamental sEMG knowledge derived from D_n is strategically adapted to the specific motions observed in D_f using advanced few-shot deep metric learning techniques. This adaptation is particularly crucial in scenarios where only a limited number of D_f samples are available, demonstrating the versatility and importance of D_n in this context. Furthermore, D_n is instrumental in the validation and testing phases of the MIA denoising methodology, as detailed in Chapter 5. Its inherent absence of motion-induced

artefacts (MIA) allows for the use of synthetic MIA, providing a unique opportunity to rigorously evaluate the denoising performance of our proposed model.



(a)



(b)

Fig. 3.3 Plots (a) and (b) display sEMG recordings from the D_f dataset, featuring different athletes and muscle groups, and illustrating the contamination of MIA on muscle contraction signals. In both plots, the x-axis represents time, and the y-axis indicates the amplitude in millivolts. Areas of MIA contamination are highlighted within red dashed rectangles. Plot (a) features two sequential MIAs during an isometric contraction of the soleus muscle. In contrast, plot (b) showcases an isotonic contraction of the biceps femoris long head muscle, also affected by MIA. These visualisations, drawn from D_f, aid in understanding the impact of MIA on sEMG signals under varying muscular activities.

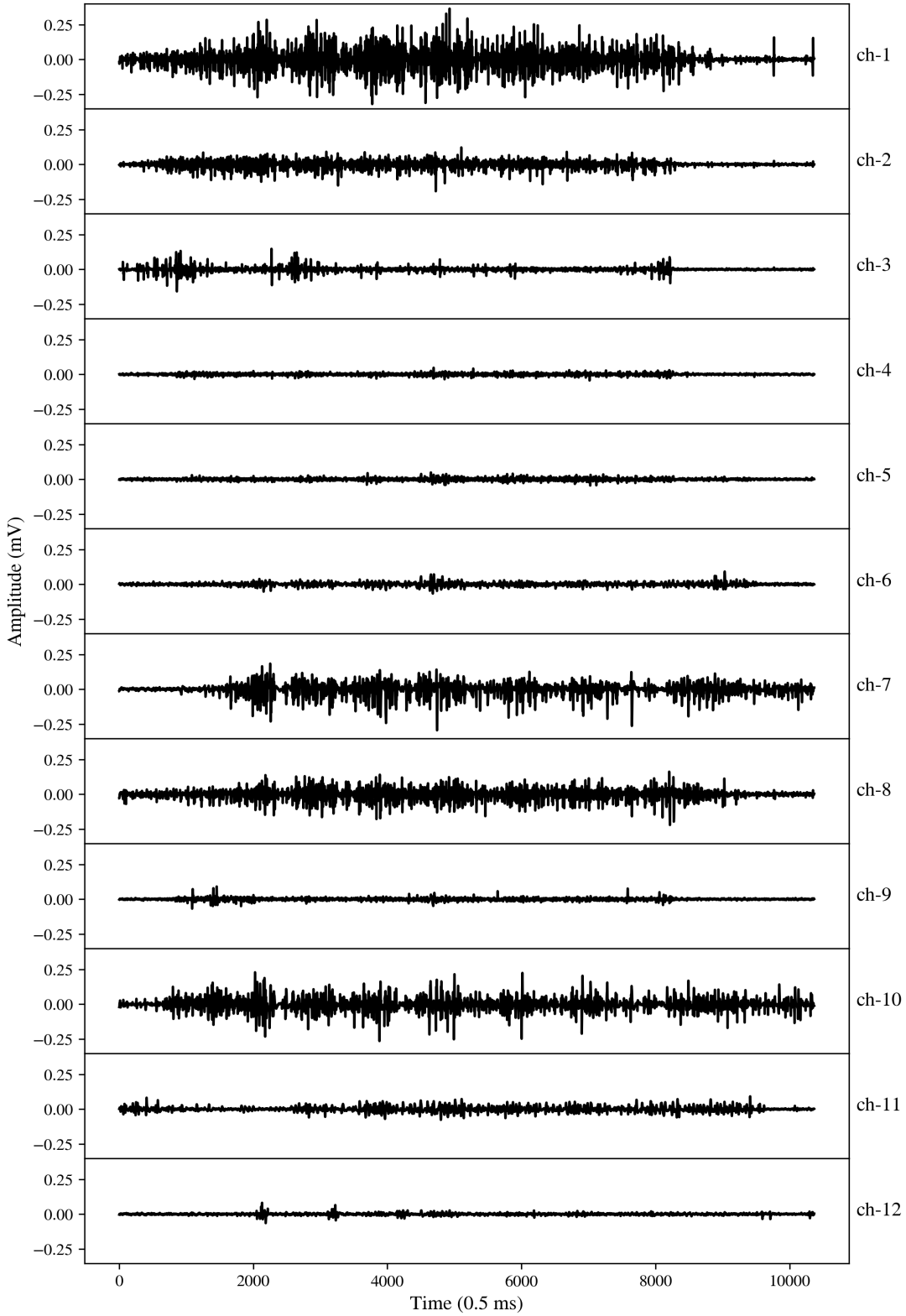


Fig. 3.4 Plot of recorded sEMG data from the knife grasping motion in the D_n dataset. The x-axis indicates time, where each index corresponds to 0.5 ms, consistent with a 2 kHz sampling rate. The y-axis measures signal amplitude in millivolts. This plot showcases 12 channels (labelled as ch), each associated with a distinct muscle. Channels 1-8 (ch-1 to ch-8) are placed around the forearm at the radio-humeral joint, capturing a range of muscle signals. Channels 9 and 10 (ch-9 and ch-10) are dedicated to the primary activity of the flexor and extensor digitorum superficialis muscles, respectively. Channels 11 and 12 (ch-11 and ch-12) target the primary activity areas of the biceps brachii and triceps brachii muscles.

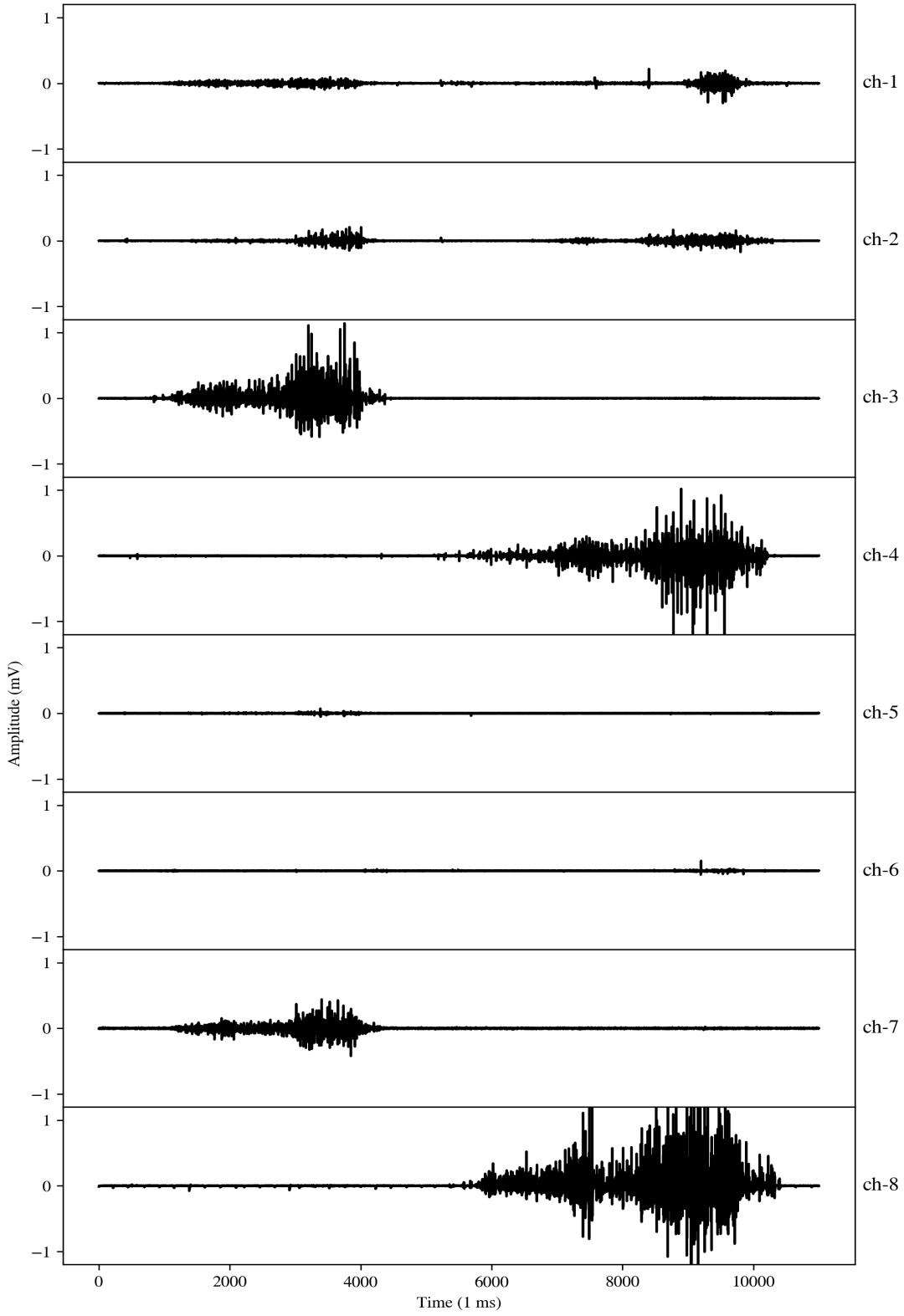


Fig. 3.5 Plot of recorded sEMG data from the hamstring claw exercise in the D_f dataset. The x-axis indicates time, with each unit representing 1 ms, corresponding to a 1 kHz sampling rate. The y-axis displays signal amplitude in millivolts. Data are collected using an 8-channel configuration (labelled as ch in the plot), tailored for lower limb muscles: Channels 1-2 (ch-1, ch-2) for adductor longus (left and right), Channels 3-4 (ch-3, ch-4) for biceps femoris long head (left and right), Channels 5-6 (ch-5, ch-6) for soleus (left and right), and Channels 7-8 (ch-7, ch-8) for semitendinosus (left and right).

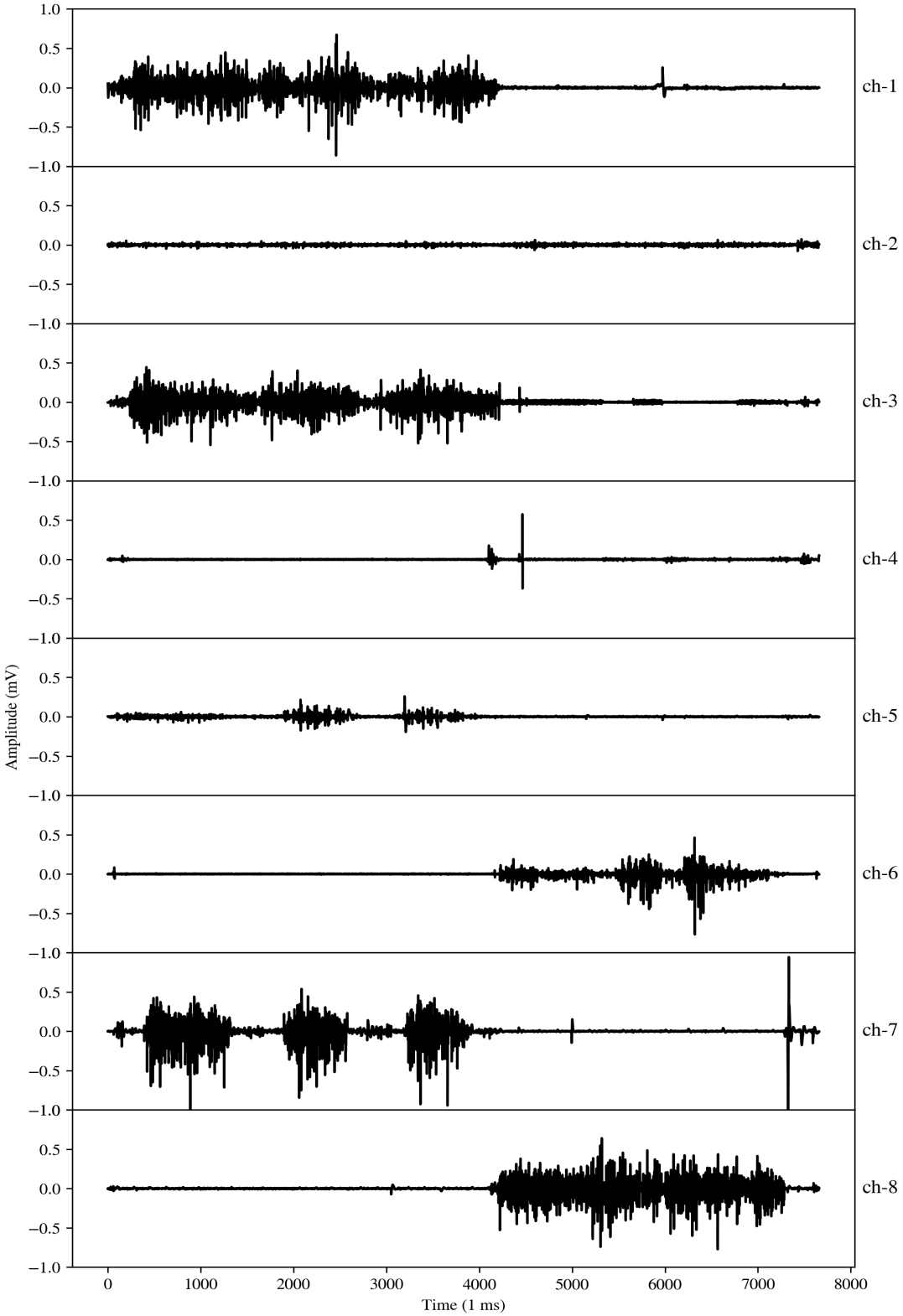


Fig. 3.6 Plot of recorded sEMG data from the SL elevated glute bridge exercise in the D_f dataset. The x-axis indicates time, with each unit representing 1ms, corresponding to a 1 kHz sampling rate. The y-axis displays signal amplitude in millivolts. Data are collected using an 8-channel configuration (labelled as ch in the plot), tailored for lower limb muscles: Channels 1-2 (ch-1, ch-2) for adductor longus (left and right), Channels 3-4 (ch-3, ch-4) for biceps femoris long head (left and right), Channels 5-6 (ch-5, ch-6) for soleus (left and right), and Channels 7-8 (ch-7, ch-8) for semitendinosus (left and right).

Chapter 4

Drill-Time Labelling

4.1 Introduction

Real-world sEMG applications face the intricate task of automating drill timing detection within sEMG recordings. Although current literature mainly employs a range of signal processing techniques to determine the onset and offset of sEMG signals, such strategies are less effective in professional sports settings [186, 114]. Traditional onset detection differentiates sEMG activity, representing muscle contraction, from inactivity deemed as noise [55]. Nevertheless, sports drills commonly involve repeated movements. In the intervals between these exercises, athletes might walk, stretch, or prepare for the subsequent drill through particular postures. Hence, the primary challenge surpasses the typical onset detection issue discussed in the literature, where the focus is primarily on differentiating contractions from noise. The real challenge lies in identifying contractions and differentiating those related to specific drills from other incidental muscle activities. The process is further complicated with the inclusion of data from 8 distinct muscle groups, deviating from the standard onset detection problems. Different drills trigger the activation of varying muscle combinations. Comprehending the correlations between these muscle activations becomes crucial to accurately identify drill timings. Although this is time-consuming, visual inspection of the signal remains the gold standard for onset detection, highlighting the complexity and the necessity for advanced automated methods in this context [79].

This chapter introduces a deep learning model, incorporating a convolutional neural network block to extract critical features from raw sEMG windows. This model also incorporates bidirectional long short-term memory recurrent neural networks (BLSTM RNN) to discern temporal patterns among the extracted feature vectors. Moreover, a novel loss regularisation, spike regularisation, is proposed to counteract frequent prediction fluctuations, considering that drills do not commonly alternate between onset and offset in quick succession.

4.2 State-of-the-art

Common onset detection methods often leverage sEMG's distinct features, such as amplitude and frequency, by contrasting them with a set threshold [164]. The linear envelope (LE) technique for onset detection is explored in [135]. Here, a low pass filter is applied to MAV of sEMG windows. While the approach is computationally efficient, it struggles to differentiate between onset and offset during low SNR or in the presence of motion-induced artefacts (MIA). Another method, adaptive contraction detection (ACD), is introduced in [114]. This frequency spectrum-based technique employs the geometric mean of the sEMG window's power spectral density for onset detection. This approach effectively deals with low SNR situations but faces challenges with MIAs since they typically span a broad frequency spectrum, leading to a raised geometric mean. The modified adaptive linear detector (M-ALED), discussed in [177], relies on the Teager Kaiser Energy (TKE) operator. This method uses the TKE median to counteract the impulsiveness of the TKE transformation, making it effective for low SNR scenarios but less so for high MIA situations. Lastly, sample entropy (SampEn), a measure of a signal's predictability, has been utilised in various applications, including onset detection. A higher SampEn value suggests increased randomness in the signal, marking it as a particularly effective metric for differentiating both onset and offset under varied conditions such as low SNR and MIA [33].

Recent studies indicate that deep learning models consistently outpace traditional signal processing techniques. In [55], a bidirectional LSTM RNN methodology surpasses all other state-of-the-art techniques when evaluated on a soccer dataset inclusive of MIA. Similarly, [42] showcases how machine learning techniques excel over established signal processing methods such as LE, RMS, and CWT.

4.3 Contributions

The main contributions of this chapter include:

- Introduction of a unique problem that bridges the gap between academic research and real-world sports applications. This problem, drill timing detection, presents added challenges to the classic onset detection task since 8-channel sEMG data is utilised, and specific drill-related activities are differentiated from unrelated ones.
- The pioneering application of deep learning for drill timing detection, conducted on data sourced from elite players across two EPL football clubs.
- Proposal of a deep learning model designed to extract essential features from each frame while capturing the temporal relationships between them using a BLSTM RNN.
- Achievement of a high accuracy rate of 96.7%, 98.4%, and 99% for subwindow sizes of 512, 1024, and 2048, respectively.

- Introduction of a novel loss regularisation technique termed spike loss regularisation, which is designed to mitigate unforeseen fluctuations in prediction data, leading to an improvement in the model's F1 score by 2.8%.

4.4 Proposed Methodology

4.4.1 Preprocessing

The preprocessing of the sEMG data begins with segmentation into discrete windows. Each window comprises eight channels, with 8192 samples per channel, resulting in a vector dimension of $(8, 8192)$. The dataset denoted as D_f, operates at a sampling rate of 1 kHz, implying that each window spans approximately 8 seconds. The primary rationale for implementing normalization on a per-window basis, rather than globally across the entire data stream, is to encourage the model to identify and leverage spectral features specific to discriminating between drill and non-drill activities.

After windowing, each window undergoes preprocessing prior to the deployment of the deep learning model. Initially, a 9^{th} order Butterworth bandpass filter is applied with cutoff frequencies set at 50Hz and 300Hz. Even though the high pass filter at a large cutoff frequency causes some valuable sEMG data loss, as sEMG has valuable data starting from 20Hz [57], the filter is determined to be the best performing for the deep learning model through various tests conducted at different cutoff frequencies. Subsequent to this, min-max normalization is executed, defined as

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{4.1}$$

This procedure is followed by rectification, which involves subtracting the mean from the normalised data and then taking its absolute value:

$$x_{\text{rectified}} = \left| x_{\text{norm}} - \mu(x_{\text{norm}}) \right|, \tag{4.2}$$

where x_{norm} denotes the min-max normalised signal of sEMG data x and $\mu(\cdot)$ represents the mean function.

Upon normalisation, every primary window, characterised by a shape of (8,8192), is further divided into subwindows, each comprising 512 samples. As a consequence, every sample, denoted as x, in the dataset D_f, adopts a shape of (16,8,512) — corresponding to the subwindow count, channel number, and sequence length, respectively. Within the dataset, for every data sample x^i, an associated true label y^i exists. This label assumes a shape of $(16, 1)$, indicating the binary true label for each subwindow. Here, i signifies the index of the (x^i, y^i) pair within the dataset D_f.

4.4.2 Network Structure

The architecture of the proposed model is depicted in Figure 4.1. The model takes an input x^i consisting of 16 subwindows, each with a dimension of $(8, 512)$, resulting in an overall shape of $(16, 8, 512)$.

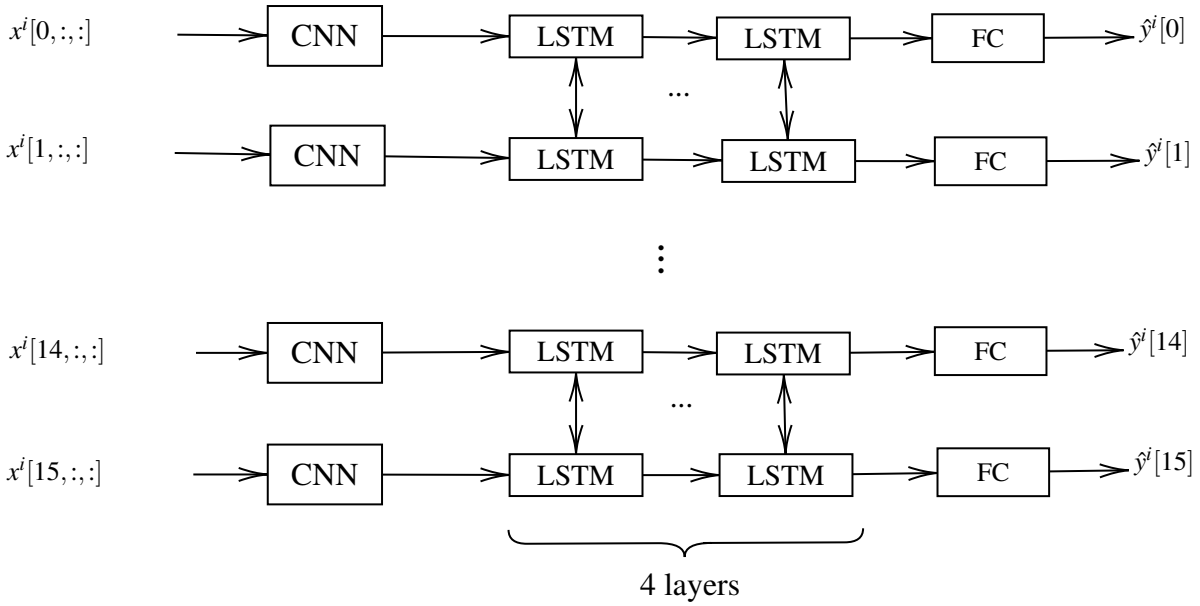


Fig. 4.1 An illustration of the network structure. The system takes an input of 16 subwindows, each with a shape of $8,512$, indicating the channel dimension and the sequence length, respectively. Each of these subwindows undergoes parallel processing in CNN blocks. Subsequently, the generated feature vectors from the CNN blocks are directed to a four-layer deep bidirectional LSTM RNN. Each LSTM instance's output is directed into a fully connected layer with a sigmoid activation function, returning a singular output. The completion of this process results in a predictive sequence with a shape of $(16,1)$, denoting a distinct prediction for every subwindow.

These dimensions correspond to the subwindow size, channel size, and sequence length, respectively. Each subwindow undergoes parallel processing in a convolutional neural network (CNN) module, with shared weights. Every CNN block outputs a feature vector of dimension 32. A detailed description of the layers, as well as the input and output shapes of the CNN block, is presented in Table 4.1. The CNN block incorporates the kernel-depth (KD) block feature extractor, shown in Figure 4.2, also proven effective for motion classification tasks on the Ninapro dataset (see Chapter 6). The KD block uses an attention mechanism to fine-tune kernel sizes and layer depth, thereby enhancing adaptability across different applications. The KD block produces a channel output and sequence length of 512. Subsequently, the KD block output is reduced to a feature vector dimension of $(32,1)$ through successive convolution, batch normalization, and adaptive pooling layers, with the ReLU activation function applied after each convolution.

The resultant feature vectors from the 16 individual subwindows are directed into the LSTM module. This module includes a 4-layer bidirectional LSTM RNN with a hidden dimension of 128 to identify the temporal relationships among the subwindows. Each LSTM output is then passed to a fully connected (FC) layer, where a multilayer perceptron with a sigmoid activation function produces a scalar output value. For an input dimension of $(16,8,512)$, the model delivers an output of $(16,1)$, representing a decision for each 16 subwindow.

Table 4.1 Layer-by-layer breakdown of the CNN model detailing layer type, input shape, and output shape for each layer. The module processes an input of shape (8,512), representing channel size and sequence length, and extracts a feature vector with a shape of (32,1).

Layer No.	Type	Input Shape	Output Shape
(1)	KD Block	(8, 512)	(512, 512)
(2)	Conv1d	(512, 512)	(256, 512)
(3)	BatchNorm1d	(256, 512)	(256, 512)
(4)	AdaptiveAvgPool1d	(256, 512)	(256, 256)
(5)	Conv1d	(256, 256)	(128, 256)
(6)	BatchNorm1d	(128, 256)	(128, 256)
(7)	AdaptiveAvgPool1d	(128, 256)	(128, 128)
(8)	Conv1d	(128, 128)	(64, 128)
(9)	BatchNorm1d	(64, 128)	(64, 128)
(10)	AdaptiveAvgPool1d	(64, 128)	(64, 64)
(11)	Conv1d	(64, 64)	(32, 64)
(12)	BatchNorm1d	(32, 64)	(32, 64)
(13)	AdaptiveAvgPool1d	(32, 64)	(32, 32)
(14)	Conv1d	(32, 32)	(32, 32)
(15)	BatchNorm1d	(32, 32)	(32, 32)
(16)	AdaptiveAvgPool1d	(32, 32)	(32, 1)

Kernel-Depth (KD) Block

The KD block is specifically engineered to enhance time series feature learning capabilities in deep neural networks. This enhancement is realised by employing 1D convolution blocks in various kernel sizes and depths. Determining the most effective model depth and kernel size often involves extensive experimentation to pinpoint the ideal parameters, which can vary based on the dataset and application.

Recent research indicates that utilising a range of kernel sizes in parallel can substantially enhance model performance [25, 86]. This improvement stems from the ability of each kernel size to capture unique characteristics of the data. In response to this finding, our model's architecture incorporates five distinct kernels within the D block framework, detailed in Section 4.4.2. Outputs from each D block and a residual tensor are merged along the channel dimension. This method of channel-wise concatenation allows the model to integrate diverse features from each kernel size. The residual tensor is produced by a 1D convolution layer with a kernel size of 1 and an output filter size of 32, ensuring compatibility in dimensions with the D block outputs.

As depicted in Figure 4.2, the model initially passes through a 1D convolution block with a kernel size of 1 to modify the channel size using an output filter size of 32. After this step, the model processes the data through the D blocks with varied kernel sizes. The data is directed to an attention block upon concatenating the outputs from these D blocks along the channel dimension. This block's role is to

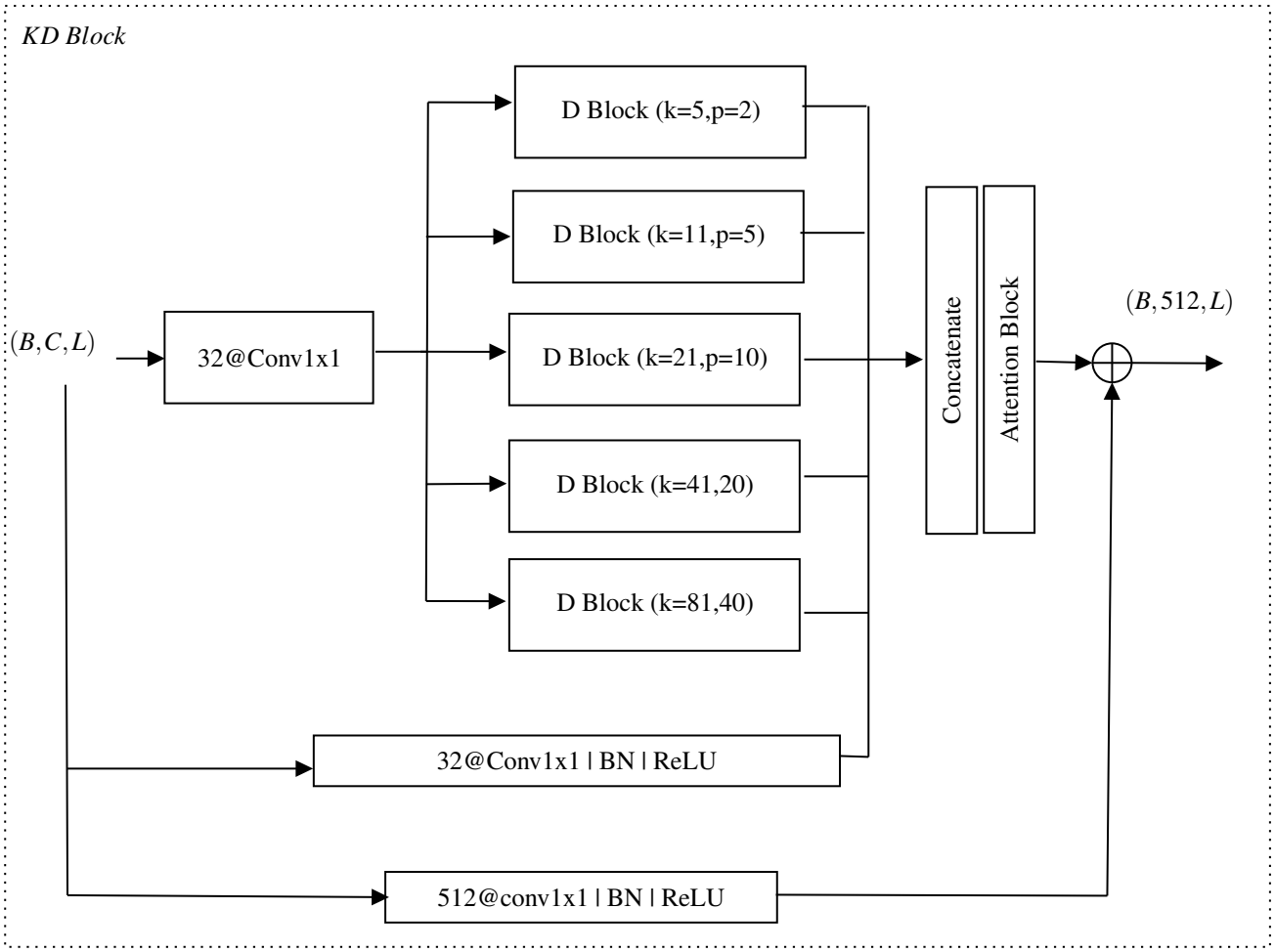


Fig. 4.2 Schematic diagrams of KD block is provided. KD block consists of an initial 1D convolution, parallel D blocks with varying kernel sizes, concatenation with a residual tensor, attention block, and summation with another residual, transforming an input tensor of shape (B,C,L) into an output tensor of shape $(B,512,L)$.

identify and prioritise the most critical kernel sizes and depths. The attention block's output is then amalgamated with a residual through element-wise addition. Incorporating this residual fulfils two essential roles: it counters the vanishing gradient issue, which can impede learning in deep networks, and it enhances the reuse of features across different layers [184]. The residual is sourced from a 1D convolution block with an output filter size equivalent to the concatenated channel count, followed by batch normalisation and ReLU activation. The KD block processes an input tensor of shape (B,C,L), representing batch size, channel count, and window length, respectively, and delivers an output tensor of shape $(B,512,L)$.

Depth (D) Block

The D block is specifically designed to explore how different depths enhance performance, especially using a series of 1D SConv blocks. Using SConv instead of traditional convolution layers brings advantages in terms of computational efficiency and reduced parameters. This approach also lowers

the risk of overfitting and improves the extraction of features, as it handles spatial and channel-based features separately [28]. Illustrated in Figure 4.3, the D block sequentially employs SConv layers, each with consistent kernel size and padding. The output from each layer is combined along the channel dimension. This design not only outputs the final SConv layer it also includes outputs from all layers to utilise residual connections. This approach prevents vanishing gradients and incorporates important features from every depth level. Each layer of the SConv contributes unique features and collectively enhances the model's performance. When these outputs are merged and processed through the KD Block's attention mechanism, the model effectively identifies depth values critical for accurate classification and then increases the focus on those key depths. The input and output filter sizes for the SConv layers in the D Block are consistently set to 32, leading to a tripling of the output channel size after the concatenation process.

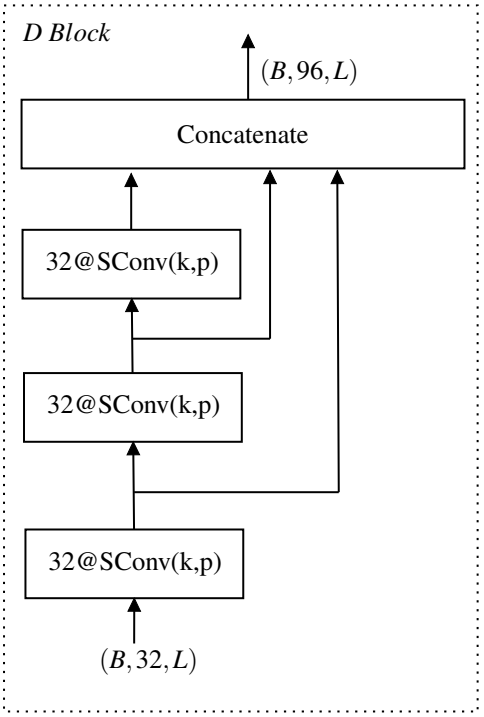


Fig. 4.3 Schematic representation of the D block provided. The D block features cascading 1D Separable Convolution blocks (SConvs) commencing from an input size of $(B, 32, L)$. After all the SConv blocks are processed, their outputs are merged, tripling the channel dimension to produce the concluding shape of $(B, 96, L)$. Combined, these structures refine feature learning by leveraging multiple depths and kernel dimensions, enhancing the overall efficacy and output of the model.

Attention Block

The attention block in our model comprises two sequential attention modules: channel attention and temporal attention. Influenced by the Convolutional Block Attention Model (CBAM) as described in [173], our method employs 1D channel and temporal attention to enhance classification accuracy in time-series data. The channel attention module is specifically tailored to identify and emphasize the most significant channels, whereas the temporal attention module concentrates on pivotal samples

across the sequence. For an input feature tensor h to the attention block, the output h' is computed as follows:

$$h' = F_t(F_{ch}(h) \otimes h) \otimes (F_{ch}(h) \otimes h), \tag{4.3}$$

where $F_{ch}(\cdot)$ and $F_t(\cdot)$ are the channel and temporal attention functions, and $\otimes$ symbolizes matrix multiplication. The channel attention function, $F_{ch}(\cdot)$, is formulated as:

$$F_{ch}(h) = \sigma(F_{mlp}(h_{\text{avg}}) + F_{mlp}(h_{\text{max}})), \tag{4.4}$$

where h_{avg} and h_{max} are the results of 1D average pooling and 1D max pooling across the temporal dimension of h. The sigmoid activation function is represented by $\sigma(\cdot)$. $F_{mlp}(\cdot)$ denotes a multi-layer perceptron comprising three FC layers. Similar to the squeeze-and-excitation network concept [81], the block uses $F_{mlp}(\cdot)$ for an enhanced attention mechanism. Initially, $F_{mlp}(\cdot)$ reduces the number of perceptrons as per a predetermined reduction ratio, before returning to the original channel number. For our KD Block, we set this reduction ratio at 16. The temporal attention function, $F_t(\cdot)$, is tasked with calculating the mean and maximum values across the channel dimension of the input tensor. This calculation yields two tensors that represent the maximum and mean values across all channels for the sequence length. These tensors are then concatenated along the channel dimension and processed through a 1D convolution block. This block, having an output filter size of 1, a kernel size of 3, and a padding of 1, concentrates on prominent temporal features while preserving spatial context. Algorithm 1 provides a detailed pseudo-code representation of how the attention module functions.

4.4.3 Spike Loss Regularisation

The model generates a prediction array $\hat{y}$, containing 16 values between 0 and 1. These values denote the probability that a subwindow is attributed to a drill or noise. Then, BCE loss for a sequence of predictions is given by

$$L_{BCES}(y,\hat{y}) = \frac{1}{N}\sum_{i=1}^{N} L_{BCE}(y[i],\hat{y}[i]), \tag{4.5}$$

where L_{BCE} is defined in Section 2.3.2.

Spike loss regularisation minimises unnecessary fluctuations in $\hat{y}$. Given that drills last approximately eight seconds, an eight-second input sequence may entirely be noise, represent a drill-noise transition, entirely be a drill, or signify a noise-drill transition. At most, two transitions are expected. Any deviation from this indicates unwanted spikes, leading to prediction errors. To compute spike loss regularization, $\hat{y}$ is rounded to 1 or 0, as shown:

$$\hat{y}'[i] = \begin{cases} 1 & \text{if } \hat{y}[i] \geq 0.5 \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

With $\hat{y}'$ being the rounded array, the discrete derivative, $\Delta\hat{y}'$, is then calculated as

$$\Delta\hat{y}'[i] = \hat{y}'[i+1] - \hat{y}'[i]. \tag{4.7}$$

Algorithm 1 1D Attention Block Module

procedure CHANNEL ATTENTION(x)
    $x_{\text{avg}} \leftarrow \text{AdaptiveAvgPool1d}(x)$
    $x_{\text{max}} \leftarrow \text{AdaptiveMaxPool1d}(x)$
    $x_{\text{avg}} \leftarrow \text{mlp}(x_{\text{avg}})$
    $x_{\text{max}} \leftarrow \text{mlp}(x_{\text{max}})$
    $x_{\text{sum}} \leftarrow x_{\text{avg}} + x_{\text{max}}$
    $x_{\text{sig}} \leftarrow \text{sigmoid}(x_{\text{sum}}) \otimes x$
    return x_{sig}
end procedure

procedure TEMPORAL ATTENTION(x)
    $x_{\text{avg}} \leftarrow \text{mean}(x, \text{dim} = 1, \text{keepdim} = \text{True})$
    $x_{\text{max}} \leftarrow \text{max}(x, \text{dim} = 1, \text{keepdim} = \text{True})$
    $x_{\text{cat}} \leftarrow \text{concatenate}(x_{\text{avg}}, x_{\text{max}}, \text{dim} = 1)$
    $x_{\text{conv}} \leftarrow \text{conv}(x_{\text{cat}})$
    $x_{\text{sig}} \leftarrow \text{sigmoid}(x_{\text{conv}}) \otimes x$
    return x_{sig}
end procedure

procedure ATTENTION BLOCK(x)
    $x \leftarrow$ Channel Attention(x)
    $x \leftarrow$ Temporal Attention(x)
    return x
end procedure

Considering a maximum of two changes in eight seconds, the spike regularization loss becomes

$$\lambda_{\text{reg}} = \gamma \times \max\left(\left(-2 + \sum_{i=1}^{15} |\Delta\hat{y}'[i]|\right), 0\right), \tag{4.8}$$

where $\gamma = 0.80$ is the regularization constant. Then, the loss function $L_{spk}(\cdot,\cdot)$ of our model becomes

$$L_{spk}(y, \hat{y}) = L_{BCES}(y, \hat{y}) + \lambda_{\text{reg}}. \tag{4.9}$$

4.5 Experiments

The dataset is partitioned into training, validation, and test sets, avoiding any overlap of drill data from a single session across these subsets to prevent data leakage. Specifically, drill data from each athlete's session included in the training set is not used in the test set. After this initial separation, the training set is shuffled. These divisions are detailed in Table 4.2. The training set is utilized for model optimisation, the validation set to mitigate overfitting, and the test set for comparing different methods. Experiments are replicated ten times using various random seeds, with results averaged to ensure robustness.

During data collection, physiotherapists annotate drill timings, marking the commencement of a drill via the Neurocess application. These annotated timings undergo verification through visual inspection of the raw data. Given the onset and offset timestamps for drills, each sEMG data sample receives a binary label: 1 for drill and 0 for noise. This study does not concentrate on sample-by-sample labeling. Instead, it assesses whether sub-windows are classified as drill or noise. Some sub-windows may comprise both drills and noise. If a sub-window encompasses 20% or more samples labelled as 1, the entire sub-window is labelled as drill; otherwise, it receives a noise label.

The signal processing methods in Section 4.1 target 1-channel data and a different problem—onset detection. Comparing these methods to our dataset is neither fair nor straightforward. Likewise, the deep learning methods from [55] and [42] are meant for a 1-dimensional onset detection problem and are unsuitable for our drill timing detection challenge. Hence, the methodology is not compared to existing onset detection methods.

The proposed spike loss regularization's effect is assessed by training and testing the model both with and without this regularization. The test without the spike loss regularization employs only BCE. Table 4.2 presents the training and testing datasets with distinct subwindow sizes: 512, 1024, and 2048. Different subwindow sizes are tested to ascertain the feature extractor CNN, as highlighted in Figure 4.1, and its efficacy in processing expansive raw data enriched with sEMG signal information.

Table 4.2 Sizes of train, validation, and test datasets corresponding to different subwindow sizes.

Subwindow Size	D_{train}	D_{val}	D_{test}
512	196224	28032	56064
1024	98119	14017	28034
2048	49056	7008	14016

4.6 Results

In Table 4.3, the spike loss regularization achieves a performance improvement over the BCE loss by 2.8%. This supports the claim that spike loss enhances model performance by penalising frequent and unnecessary prediction deviations. Given that athletes typically do not transition between drills and non-drills more than twice within 8 seconds, the spike loss regularization proves effective by penalising undesired spikes.

Furthermore, the model with spike loss regularization is evaluated at various subwindow sizes: 512, 1024, and 2048. As observed from Table 4.3, the F1-score tends to be directly proportional to the subwindow size. Specifically, with increasing subwindow size, the F1 scores also rise, achieving values of 96.7, 98.4, and 99 for sizes 512, 1024, and 2048, respectively. This suggests that larger subwindow sizes allow the CNN feature extractor to yield more distinct feature vectors, thereby enabling the LSTM and FC modules to classify drill timings with heightened accuracy. However, there exists a trade-off: enlarging the subwindow size increases the potential for sEMG data loss. In instances of false negatives, a more substantial subwindow size leads to losing more sEMG data.

Table 4.3 F1 scores achieved by the proposed model when comparing the effects of spike loss and BCE loss across different subwindow sizes.

Subwindow Size	Spike Loss	BCE Loss
512	96.7	93.9
1024	98.4	-
2048	99.0	-

Moreover, the incorrect predictions by the models predominantly occur not from misclassifications of entire drills but during the transitions between non-drill and drill activities, specifically at the beginnings and ends of drills. Figure 4.4 and Figure 4.5 illustrate sEMG data plots with true drill labels and model predictions, demonstrating that most misclassifications occur at these critical transition points. Additionally, Figure 4.5 highlights the performance advantage of spike loss regularization over binary cross-entropy (BCE) loss. The model using spike loss regularization reliably predicts the drill labels for the central segments of the drill, whereas the BCE loss model inaccurately identifies a subwindow at approximately the 16^{th} second as non-drill activity, leading to a false negative. This superiority of spike loss regularization is attributable to its constraint on the model to limit transitions to less than 3 within an 8-second sEMG window, as exemplified in the plot shown in Figure 4.5, which underscores its enhancement of model performance.

4.7 Conclusion

In the context of modern sports applications, especially given the rise of data-driven analytics, this chapter bridges a significant gap. It extends academic paradigms to pragmatic, real-world sports scenarios. The novel problem of drill timing detection, especially with the use of 8-channel sEMG data, stands distinct and presents challenges differing from conventional onset detection tasks.

The introduced deep learning model is tailored for drill timing detection, utilising data from elite athletes in EPL football clubs. This underscores both the practicality and the significant relevance of the research in elite sports performance analytics. The employment of a BLSTM emphasizes the indispensable role of temporal relationships between frames for precise drill detection. The application of KD block is effective as a feature extractor for sEMG raw data across various scenarios. Increasing the kernel size and selecting optimal depth enhances the versatility and performance of the CNN-based feature extractor.

The spike-regularisation loss technique notably enhances the model's F1 score by 2.8%. This enhancement suggests the potential utility of this method in other prediction tasks. Additionally, the performance alterations with different subwindow sizes infer that it's possible to fine-tune the model based on specific data characteristics. However, it also underscores the challenge of striking a balance between accuracy and potential data loss.

Furthermore, as illustrated by Figure 4.4 and Figure 4.5, the proposed model tends to generate false predictions predominantly at the edges of the windows, rather than at the centers or across

entire windows. This misclassification behavior is likely due to reduced information availability at the window's edges compared to its center. The bidirectional RNN with LSTM architecture effectively channels more comprehensive information across the central sub-windows, facilitated by data flow from both directions, whereas the edges receive less such integration. Additionally, variations in true labeling may contribute to the higher error rate at the window edges. Labeling inconsistencies arise because initial labels are applied by physiotherapists in real-time during exercises and subsequently refined upon expert review of offline data. This process might introduce slight variations in the start and stop indices of drills, further contributing to misclassifications at these critical points. As a potential direction for future work, implementing majority voting with sliding overlapping windows could be considered. This approach ensures that the edges of the windows are continuously processed and interpreted as central and mid-points. Employing majority voting might improve accuracy and address the issue of edge misclassifications, thereby enhancing model reliability in identifying true drill activities.

This chapter addresses the challenge of drill timing detection in sports using a specialized deep learning model. Particularly for data originating from elite English Premier League (EPL) clubs, the approach achieves accuracies of 96.7%, 98.4%, and 99% for window sizes of 512, 1024, and 2048 respectively. Incorporating the spike-regularization technique, the model improves the binary cross-entropy (BCE) loss by 2.8%. Deep learning exemplifies its potency for drill timing detection in sports contexts. Despite the success of the proposed methodology in determining the start and end of drills in the sports domain, it is not suitable for real-time or online applications, such as real-time hand gesture recognition, where decisions about the onset need to be made within 0.12 s [60]. The proposed methodology requires at least 8 seconds long windows and makes decisions afterwards, which renders it more appropriate for offline drill-timing detection applications. Additionally, the sub-window size, which is 0.5 s in this chapter, is considered too long for hand gesture recognition use cases, where smaller window sizes are needed for precise onset labeling of various muscles. Thus, it can be concluded that the methodology described in this chapter is designed primarily for lower extremity drill-time labeling for sports exercises.

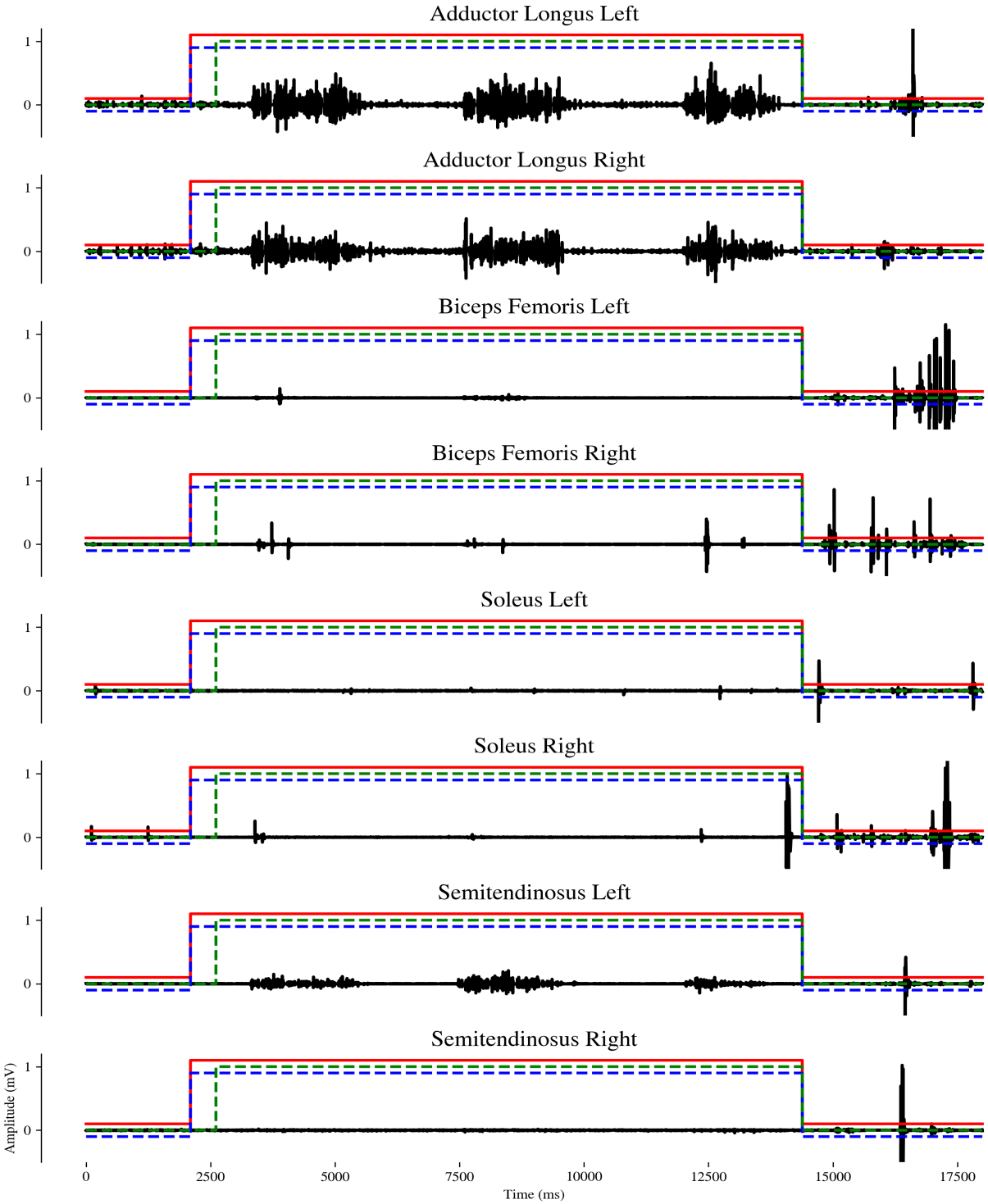


Fig. 4.4 Plot of sEMG data (black line) from the D_f dataset, showing true drill labels (red line), predictions from a model trained with spike loss regularization (blue dashed line), and predictions from a model trained with binary cross-entropy loss (green dashed line). The x-axis represents time, with each index corresponding to a 1 ms interval, consistent with a 1 kHz sampling rate. The y-axis measures the signal amplitude in millivolts. The model utilizing spike loss regularization accurately predicts the sequence of true drill labels, whereas the model trained with binary cross-entropy loss fails to identify the initial drill label window.

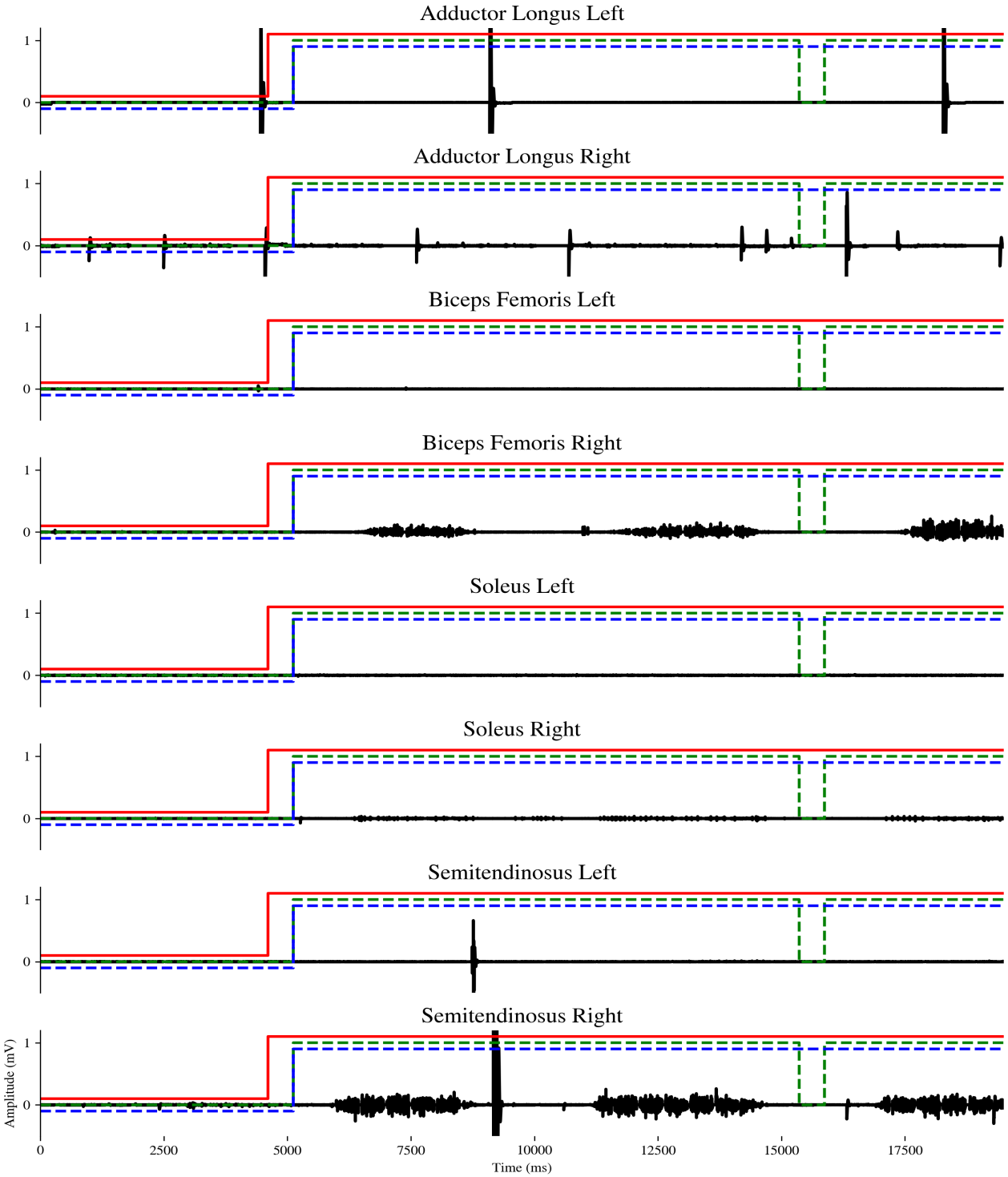


Fig. 4.5 Plot of sEMG data (black line) from the D_f dataset, illustrating true drill labels (red line), predictions from a model trained with spike loss regularization (blue dashed line), and predictions from a model trained with binary cross-entropy (BCE) loss (green dashed line). The x-axis represents time, with each index corresponding to a 1 ms interval, consistent with a 1 kHz sampling rate. The y-axis measures the signal amplitude in millivolts. Both models fail to detect the initial drill label subwindow. The spike loss regularization model accurately predicts subsequent drill labels, whereas the BCE loss model incorrectly interprets a subwindow at approximately the 16^{th} second as non-drill activity, resulting in a false negative.

Chapter 5

Motion-Induced Artefact (MIA) Detection and Denoising

5.1 Introduction

In professional sports, athletes' movements and collisions produce additive noise in sEMG recordings, known as Motion-Induced Artefact (MIA) [142]. This interference makes processing sEMG signals more difficult, adversely affecting onset detection, motion classification, and feature extraction [55]. Filtering out MIA from the sEMG signal is challenging, given that sEMG and MIA share overlapping frequency spectra, exhibit similar non-stationary characteristics, and possess varied temporal shapes [4, 150, 142]. Moreover, the impulsive nature of MIA, with its unpredictable occurrence times, poses challenges for traditional noise filtering techniques [55].

In response, this section divides the problem of MIA cancellation into two subproblems: determining MIA zones within the sEMG stream and reconstructing the clean sEMG in the affected zones. It is shown that using separate models for detecting and cancelling out MIA proves to be more effective in noise cancellation than employing a singular model for both tasks.

This chapter employs two datasets to assess the model's efficacy: D_n and D_f. In this chapter, a mathematical model to synthetically generate MIA is introduced to enable the addition of artificial MIA to D_n, which is originally free of MIA as detailed in Chapter 3. The model's detection accuracy is also verified with D_f dataset, which contains many naturally occurring MIA. This experimental approach facilitates the validation of the model against both synthetic and naturally occurring MIA data.

A novel method aimed at denoising sEMG signals from MIA, a U-net Encoder-Decoder (UED) structure is proposed. The UED structure comprises unique convolutional blocks that employ attention-enhanced kernel and channel selection, proving efficient for denoising due to its capability to reconstruct high-resolution features and precisely pinpoint details, thereby facilitating the effective separation and removal of noise from the original signals.

5.2 State-of-the-art

Research on cancelling MIA in sEMG signals is limited, but extensive literature exists on methodologies for removing such artefacts in electrocardiogram (ECG) signals [90, 139]. In [139, 138, 159], an adaptive filtering approach is employed to remove the MIA component from ECG signals. This approach, though computationally efficient, requires a noise reference and can fail if there are changes in the noise signal [90]. A wavelet-based approach is explored in [178] to cancel out noise components, and [104] proposes an empirical mode decomposition-based method to generate clean signals. However, all these methods are surpassed by the deep neural network (DNN)-based Denoising Auto-Encoder (DAE) model presented in [176], which is further refined by a fully convolutional network DAE in [27]. Currently, the model proposed by Jhang et al., the channel-wise average pooling and 1D pixel-shuffle DAE (CPDAE), which integrates skip connections and pixel shuffling into the autoencoder, stands as the best-performing model in the literature for denoising ECG signals of MIA [90]. The research identifies the aforementioned methodologies used in ECG as the benchmarks for MIA cancellation and adapts the top-performing model, CPDAE from [90], to fit the sEMG signals and datasets under analysis. CPDAE is used as a reference method to compare the proposed methodologies.

5.3 Contributions

The contributions of this chapter are as follows:

- This chapter pioneers the testing of MIA detection on sEMG data using deep learning techniques. The proposed methodology is rigorously evaluated on two diverse datasets: one enriched with synthetic MIA on D_n, and the other embodying real MIA recorded on D_f.
- This chapter confirms that employing two separate models—one for detecting the MIA zone and the other for denoising—is more effective for MIA denoising in sEMG. This framework attains an SNR improvement of 17.20 dB, elevating the best performing state-of-the-art model by 7.01 dB.
- The proposed method for MIA detection attains macro and micro F1 scores of 95% and 99%, respectively, on D_f, and average macro and micro F1 scores of 94.8% and 98.8%, respectively, on the D_n with added synthetic MIA. This showcases the method's effectiveness under diverse conditions, encompassing both academic research and practical, real-world applications.
- A novel convolution block featuring attention-enhanced kernel and channel selection is utilized within the UED model. This block utilizes various kernel sizes to select the most suitable channel-kernel combination via an attention module, leading to superior performance in distinguishing noise features as different kernel sizes capture unique spatial features. The attention mechanism improves the model SNR by 0.26 dB.
- This chapter confirms that the application of MIA detection via BLSTM enhances the denoising performance not only of the proposed UED but also of state-of-the-art methods. When incorpo-

rated within the proposed framework, the denoising SNR improves by 3.12 dB, 5.48 dB, and 1.48 dB for UED, [90], and [39], respectively.

5.4 Problem Definition

The primary focus is on the denoising of the sEMG signal from the interference of MIA. In the relevant literature, motion artefact is a term often attributed to the additional noise induced by muscle movements during contractions. As highlighted in [39], this specific type of noise is generally below the 20 Hz band and can typically be filtered using a high-pass filter with a 20 Hz corner frequency and a slope of 12 dB/oct. However, this research addresses a variant of MIA predominantly arising due to two factors: physical contact with the sensor or body during exercises and the impact of the body connecting with the ground in activities such as running or jumping. As illustrated in Figure 5.1, the spectral density of the recorded MIA and sEMG during the exercises of EPL players are overlapping. Using any filter to remove the MIA band could also block valuable sEMG signals and result in the loss of necessary sEMG data.

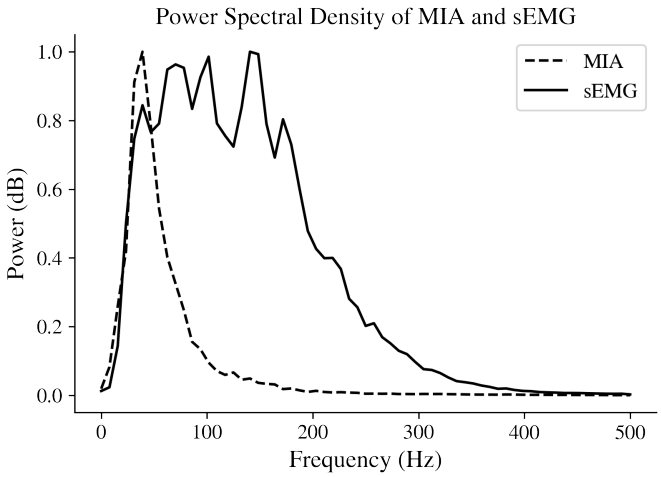


Fig. 5.1 Illustrations of normalised power spectral density (PSD) for sEMG signal, denoted as $x[n]$, and MIA, represented as $q[n]$. These plots are generated based on the sEMG data and recorded noise obtained during the experiments.

5.5 Proposed Methodology for MIA Detection

BLSTM RNN is employed to identify the presence of MIA within given sEMG streams, thanks to its ability to capture temporal patterns and dependencies in sequential data. The proposed model ingests preprocessed sEMG data consisting of 1024 samples and consists of two BLSTM RNN hidden layers, each with a hidden size of 128 and time steps of 1024. The network outputs 1024 values, each processed by separate fully connected layers with sigmoid activation function, resulting in 1024 corresponding predictions. Each prediction reflects the probability of MIA presence at a specific sEMG sample.

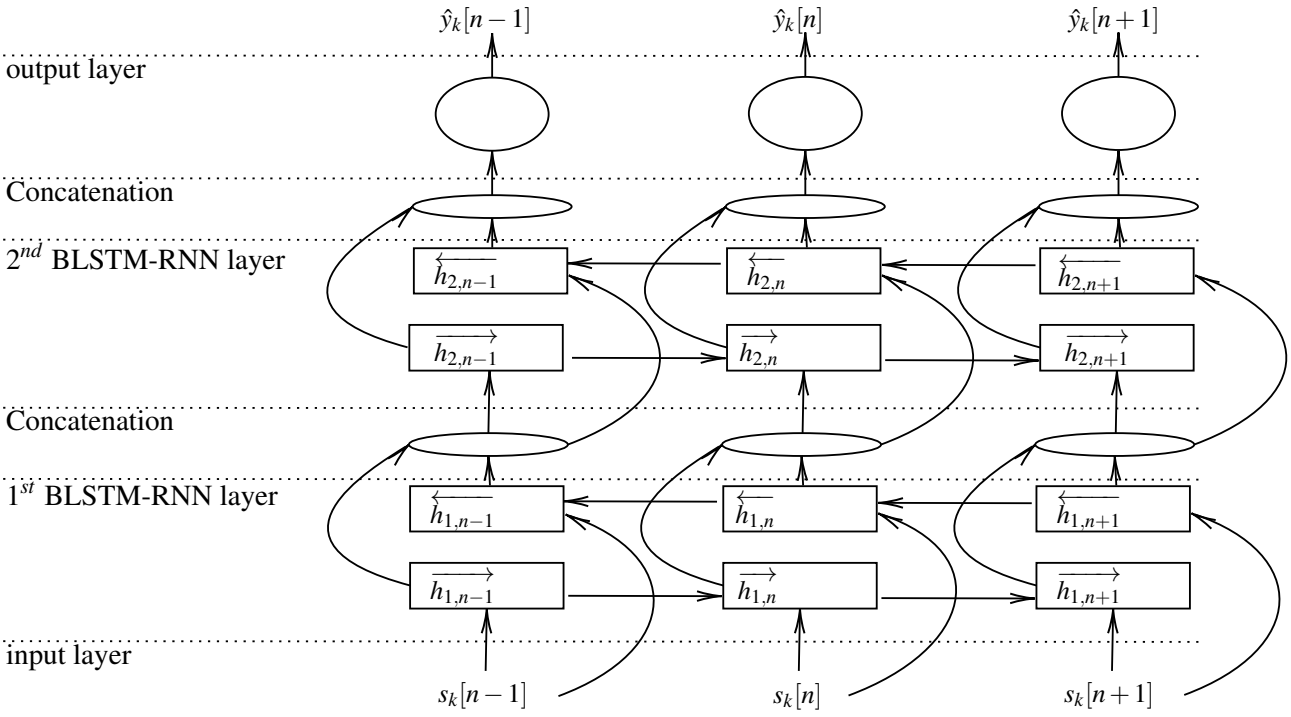


Fig. 5.2 Unfolded architecture of the proposed BLSTM-RNN model. The model processes raw sEMG data sequences, with each sEMG sample corresponding to an LSTM block. This ensures that the number of RNN time steps matches the total number of input sEMG samples. Both the forward layer vector $\overrightarrow{h}$ and the backward layer vector $\overleftarrow{h}$ are concatenated after the first and second hidden layers. Each hidden vector at a specific time step contains a feature vector with a length of 128. After concatenation, the resulting feature vectors, now of length 256, are input to a fully connected (FC) layer with a sigmoid activation function. This returns the probability that the corresponding sEMG sample belongs to a MIA.

The bidirectionality of the BLSTM RNN plays a crucial role in enhancing the model's capability. It allows the network to analyse sEMG sequences both forwards and backwards, providing a comprehensive understanding of each sample within its broader temporal context. This dual-direction analysis significantly improves the model's ability to discern whether an sEMG sample is part of an MIA or a genuine muscle signal. By considering temporal information from both directions, the model effectively captures subtle nuances and transitions within the sEMG data, leading to a more accurate identification of MIA.

The loss function employed is the Binary Cross-Entropy (BCE) complemented by spike loss regularisation, as introduced in Chapter 4. Spike loss regularisation aims to minimise superfluous fluctuations in $\hat{y}$. Given a window of 1024 samples, which typically includes one MIA, this corresponds to a maximum of two transitions in $\hat{y}$. Additionally, the ADAM optimiser is utilised for the training process.

5.6 Onset Detection with MIA in sEMG Streams

To evaluate the proposed BLSTM-RNN model, as detailed in Section 5.5, we focus on its capacity to differentiate MIA patterns from sEMG data through an onset detection task, a widely studied area in the literature. This enables a direct comparison of our model's capabilities against established methods. Given the sparse research on MIA detection, our primary goal is to achieve precise detection of muscle activity within MIA-contaminated sEMG streams. Consequently, we have specifically chosen the task of onset detection to align our evaluations with the broader scope of documented studies, facilitating a meaningful assessment against state-of-the-art approaches.

For a focused analysis, we employ a specific dataset that includes natural MIA occurrences alongside sEMG activity, as detailed in Section 5.6.2. Notably, most onset detection methods, as referenced in Section 5.6.1, utilise smaller sEMG data windows for classification, typically smaller than the 1024 sample windows described in Section 5.5. Therefore, for this experiment, the BLSTM-RNN model processes sEMG windows of 128 samples to ensure fair comparison with established methods. Accordingly, the number of RNN time steps in our BLSTM-RNN model is adjusted to 128 for this experiment. However, in the subsequent section, Section 5.7, which focuses on MIA denoising, we utilize a window size of 1024 samples, as previously outlined in Section 5.5.

5.6.1 State-of-the art onset detection methods

The evaluation of our MIA detection model is divided into two main parts. The initial part, covered in Section 5.6, concentrates on assessing the BLSTM RNN model's ability to pinpoint the onset of muscle activity in sEMG streams affected by MIA. This aspect of the evaluation is particularly important, as it allows us to compare the capabilities of our model against established methods in this area. Due to the relatively sparse research on MIA detection in the existing literature, our emphasis lies in the precise detection of muscle activity within sEMG streams contaminated by MIA.

Contemporary methods extract a scalar feature value h_k from the k^{th} window $s_k[n]$. The classification decision is ascertained by contrasting this feature with a threshold value τ_k, as illustrated in:

$$\begin{cases} \mathcal{H}_0 : h_k \leq \tau_k & \text{(offset)} \\ \mathcal{H}_1 : h_k > \tau_k & \text{(onset)} \end{cases} \tag{5.1}$$

This section compares the proposed MIA detection model in Section 5.5 with state of the art onset detection methods which are given as follows:

- **Envelope Method:** The sEMG envelope method is the most commonly used approach for quasi-online sEMG detection due to its simplicity and low computational cost [88]. First, the mean absolute value (MAV) of window $s_k[n]$ is calculated. MAV returns the mean of the fully rectified signal. To obtain the feature h_k in equation 5.1, the MAV values are smoothed by a lowpass filter with a cutoff frequency of 50 Hz [88].

- **Sample Entropy Method (SampEn):** To define SampEn, consider the Chebyshev distance between any two vectors p and q as

$$D_{\text{Chebyshev}}(p,q) := \max_i |p_i - q_i|. \tag{5.2}$$

SampEn first separates window $s_k[n]$ into windows of size w using a sliding window technique, giving

$$S_k^w[i] = [s_k[i+k]]_{k=0}^{w-1}. \tag{5.3}$$

Then the parameters N^{w+1} and N^w are defined as

N^{w+1} is the number of windows of length $w+1$ in window $s_k[n]$ having $d[S_k^{w+1}[i], S_k^{w+1}[j]] < r$,

N^w is the number of windows of length w in window $s_k[n]$ having $d[S_k^w[i], S_k^w[j]] < r$.

Typical values are $w = 2$, and r is calculated as 0.25 times the standard deviation of the sEMG window $s_k[n]$ [186]. The Sample Entropy of $s_k[n]$ is:

$$SampEn(s_k) = -\log\left(\frac{N^{w+1}}{N^w}\right). \tag{5.4}$$

- **Modified Adaptive Linear Energy Detector Method (M-ALED):** In M-ALED, the Teager-Keiser Energy (TKE) operator is applied to window $s_k[n]$ to increase the signal-to-noise ratio (SNR) by evaluating the instantaneous frequency and amplitude variations of sEMG signals, which is given by

$$\varphi_k(n) = s_k^2[n] - s_k[n+1]s_k[n-1], \quad for \quad n = 2,...,N-1. \tag{5.5}$$

TKE increases the impulsiveness of the signal. Let φ_k' be the sequence obtained by ordering $\varphi_k(n)$ in ascending order. Then, the M-ALED metric is defined as

$$\text{MALED}(s_k) = \begin{cases} \varphi_k'\left(\frac{N}{2}\right) & \text{if } N \text{ is even} \\ \frac{\varphi_k'\left(\frac{N-1}{2}\right) + \varphi_k'\left(\frac{N+1}{2}\right)}{2} & \text{if } N \text{ is odd} \end{cases} \tag{5.6}$$

- **Adaptive Contraction Detection Method (ACD):** ACD employs a frequency spectrum-based technique to detect sEMG signals. Given an sEMG window $s_k[n]$, we denote its power spectral density as P_k. The ACD metric is then computed using a geometric mean to average the power values and subsequently applying a logarithmic transformation to mitigate impulsive spectral interference. Formally, ACD is expressed as:

$$ACD(s_k) = \log\left(\prod_{i=1}^{K}(P_k[i]+1)^{\frac{1}{K}}\right), \tag{5.7}$$

where K is commonly set to $K = 128$.

The threshold update rule for the aforementioned state of the art methods is as follows. Each window $s_k[n]$ is sequentially examined using the decision in (5.1), adopting a quasi-online onset detection strategy. A salient feature of this process is the adaptability of the threshold. For accurate classification in (5.1), it's essential that the threshold surpasses the noise window's feature value. Acknowledging the potential temporal variations in noise power, a quasi-online threshold update method inspired by [14] is incorporated. Only for windows identified as noise, the threshold τ_k updated. This update is based on a linear combination of the running mean $\overline{\mu}_k$ and variance σ_k^2 of the current noise window's feature value:

$$\tau_k = \lambda_1 \overline{\mu}_k + \lambda_2 \sigma_k^2, \tag{5.8}$$

where $\lambda_1, \lambda_2 \geq 1$. The threshold update employs Welford's method, ensuring computational efficiency. Welford's approach provides a robust way to compute a running mean and variance, avoiding direct summation or squaring of large numbers which is particularly beneficial for streaming data [168]. Optimal values for λ_1 and λ_2 are determined individually for each metric and are presented in Table 5.1.

Table 5.1 Optimized λ_1 and λ_2 values per metric.

	Envelope	**SampEn**	**M-ALED**	**ACD**
λ_1	1.7	2	7	6
λ_2	2	1	2	2

5.6.2 Experiments and Results

To evaluate the performance of the proposed BLSTM RNN method for onset detection, it is benchmarked against the state-of-the-art methods detailed in Section 5.6.1. Specific to this experiment, sEMG data is collected from five Leeds United Football Club (LUFC) first-team athletes. These exercises induce impact and movement in muscles and sensors, leading to a MIA.

Four sEMG sensors are positioned on the biceps femoris and semitendinosus bilaterally as described in Chapter 3. The athletes perform various hamstring exercises: 90° isometric bridge, glute bridge, hamstring prone curls, deadlifts, squats, hamstring sprinter exercise, sprints, and non-counter movement jumps. Descriptions of notable exercises include the glute bridge, executed by elevating the hips when lying supine with knees bent at 90°. Its isometric variation involves holding the raised position for 8 seconds. The hamstring sprinter is a single-leg glute bridge with a 45 cm elevation, while the non-counter movement jump is a type of vertical leap executed from a seated position. Squats involve lowering the body by bending the knees and pushing the hips back, ensuring the back is straight and the chest up. Deadlifts entail lifting a loaded barbell off the ground to the hips and then lowering it back down. The movement is driven primarily by extending the hips and knees, ensuring the spine remains neutral throughout.

During data acquisition, each athlete undertakes three sessions, each lasting around 20 minutes on average. Given data from five athletes and four sensors in every session, the dataset encompasses 1226 minutes of sEMG recordings. The LUFC staff collaborated to label this data. LUFC physiotherapists onsite annotate the beginning (onset) and end (offset) of each training segment using sEMG software. Subsequently, these labels undergo manual verification.

sEMG data streams are segmented into windows, which receive categorisation: windows showcasing an sEMG signal acquire the 'onset' label, whilst those devoid of sEMG activity receive the 'offset' label. It is pivotal to highlight that the experiment's objective revolves around discerning noise windows, absent of sEMG signals but abundant in MIA interference, from genuine sEMG signals. Since the experiments include real-life exercises realised by atheltes, instances of MIA interference blending with sEMG signals do arise. Yet, such windows inherit the 'onset' label, attributed to their encapsulation of muscle contraction. This chapter later elaborates on the extraction of MIA from active sEMG onset regions.

In this experiment, data from one subject is used for training the BLSTM RNN model, and data from the remaining four subjects are used for testing. This design deviates from the traditional Leave-One-Subject-Out (LOSO) method, where each subject is sequentially left out for testing while the others are used for training. Our approach, which involves using data from a single subject for training and the remaining subjects for testing, facilitates a larger and more comprehensive test set. This setup is predicated on the assumption that the MIA variables are predominantly influenced by the type of activity and impact rather than by individual subject characteristics. This assumption is crucial as it directs the design and interpretation of our experimental results. The training dataset is randomized after the train-test split to prevent the model from learning sample relationships that could bias its generalization capabilities. Conversely, the test data retains its chronological order, allowing for a direct comparison between the state-of-the-art signal processing methods in Section 5.6.1 and the proposed model. The specifics of the datasets, including total windows and label distribution, are summarized in Table 5.2.

Table 5.2 Overview of training, validation, and test datasets including total windows, samples, duration, and label distributions.

	Total windows	**Total samples**	**Total minutes**	**Onset %**	**Offset %**
Training Set	131624	4212064	70.20	47 %	53 %
Validation Set	14626	468128	7.80	50 %	50 %
Test Set	2152960	68894816	1148.24	36 %	63 %

Table 5.3 Accuracy, F1 score, sensitivity, specificity, and precision metrics for the state-of-the-art methods and proposed method.

	Envelope	**SampEn**	**M-ALED**	**ACD**	**BLSTM-RNN**
Accuracy	61.41	84.95	58.86	65.54	96.73
F1-Score	58.10	78.06	41.24	64.53	95.51
Sensitivity	73.67	73.72	39.74	86.29	95.82
Specificity	54.41	91.35	69.76	53.70	97.24
Precision	47.96	82.94	42.84	51.52	95.20

Table 5.4 Confusion matrices (in percentages) for onset and offset states during experiments across all the mentioned techniques. The percentages are derived based on the entire sample count, offering a streamlined view of distribution.

	Envelope		**Sampen**		**M-ALED**		**ACD**		**BLSTM-RNN**	
	0	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
0	27%	29%	58%	5%	44%	19%	31%	30%	62%	2%
1	9%	35%	10%	27%	22%	15%	5%	34%	1%	35%

In the experiments, the methods from Section 5.6.1 underperformes compared to BLSTM RNN. As depicted in Table 5.3 and Table 5.4, BLSTM RNN attained the highest accuracy of 96.73% and an F1 score of 95.51%. BLSTM RNN accuracy surpasses the next best, SampEn, by 11.78%. SampEn, although second, is noted for its resilience against MIA interference [186].

Figure 5.3 demonstrates the difficulties encountered by state-of-the-art methods in handling MIA windows, often resulting in false positives. This issue largely stems from their dependence on spectral or time-domain features of sEMG signals, which struggle with the overlapping frequency spectra and similar non-stationary characteristics of MIA and sEMG signals [4, 150, 142]. It is important to note that Figure 5.3 is not derived from the D_f dataset but rather from a dataset specifically collected for this experiment, as previously mentioned. The figure shows MIA's dominant and high-amplitude impact between drills, a challenge for most state-of-the-art signal processing methods that fail to distinguish MIA-contaminated signals in offset regions from onset zones. The true labels in the figure, marked in red, were initially annotated by the Leeds United staff. Notably, the SampEN method performs second best, identifying some MIA zones as offset. The most significant indicator of the BLSTM RNN model's performance in this figure is its dual capability: detecting low sEMG activity as onset within noisy baselines and accurately identifying high amplitude MIA offset zones, even when these are close to sEMG power levels. The key advantage of our model lies in its use of temporal patterns rather than power-based features to discriminate between MIA and actual sEMG waveforms. This approach

enables more accurate differentiation, underscoring the proposed model's effectiveness in handling complex MIA scenarios within sEMG data.

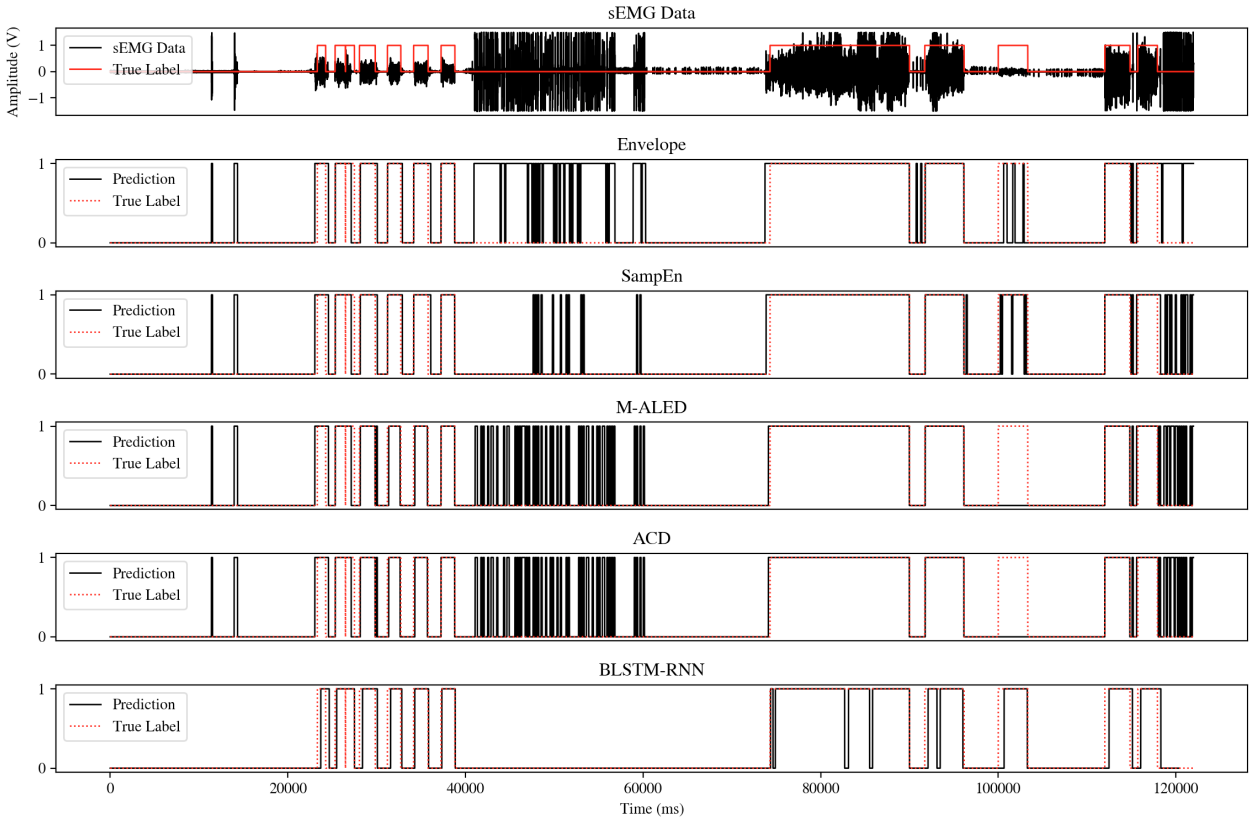


Fig. 5.3 The illustration of an sEMG recording from the experiments, alongside the true onset/offset labels and predicted labels by various methods: Envelope, SampEn, M-ALED, ACD, and BLSTM RNN. Notably, the high MIA interference observed between the 40^{th} and 60^{th} seconds could be attributed to sensor flapping caused by weakened attachment on the body, a rare occurrence typically following high-intensity exercises. This happens when medical tapes lose their adhesiveness, leading to sensors flapping against the body. The depicted data segment specifically aims to showcase the BLSTM RNN's precise performance under such extreme conditions.

5.7 Identifying MIA-Contaminated Zones in sEMG Regions

This section evaluates the effectiveness of the proposed BLSTM RNN model in identifying MIA-contaminated areas within active sEMG signals. We utilize two distinct datasets for this validation: D_n and D_f. Dataset D_f comprises naturally occurring MIA instances captured during data collection, while D_n includes synthetic MIA, intentionally introduced to facilitate controlled variation in MIA magnitude. This experimental setup of synthetic MIA is designed to assess the model's performance under conditions where MIA and sEMG signal amplitudes are comparable, thus simulating a challenging detection environment.

5.7.1 Synthetic MIA Model

Let $\hat{x}[n]$ represent the sEMG signal devoid of MIA and $q[n]$ represent the MIA. The recorded sEMG signal stream, $x[n]$, can thus be expressed under the assumption that MIA behaves as additive noise:

$$x[n] = \hat{x}[n] + q[n] \tag{5.9}$$

This model crucially assumes the MIA to be an additive noise to the sEMG data, which simplifies the analysis but does not capture more complex interactions such as changes in electrode position relative to the EMG source, which could introduce multiplicative noise characteristics. Nonetheless, $\hat{x}[n]$ signifies the sEMG signal obtained without MIA, inherently containing white Gaussian noise and 50 Hz power line interference. However, the central aim of this research is isolating and addressing the impact of MIA; consequently, the objective is to accurately retrieve $\hat{x}[n]$ from $x[n]$. The signal $\hat{x}[n]$ is not synthetically generated but is mainly derived from the Ninapro dataset, comprising hand gesture recordings where no MIA interference is present. This analysis explicitly highlights that the assumed additive nature of MIA in this model is a simplification, serving primarily to facilitate the experimental investigation of noise removal techniques.

The MIA, $q[n]$, in (5.9) is constructed based on [55], where it is presumed that body movement induces a disturbance depicted by a sequence, $\gamma[n]$. Thus, $q[n]$ is represented as

$$q[n] = p[n] * \gamma[n] = \sum_i \gamma[n-i], \tag{5.10}$$

where $p[n]$ is a series of discrete impulses denoted as $p[n] = \sum_i \delta[n-i]$, with i indicating the instances of impact caused by body movements. The sequence $\gamma[n]$ is delineated as a linear combination of three orthonormal basis sequences as

$$\gamma[n] = \sum_{i=0}^{2} a_i \phi_i[n], \tag{5.11}$$

and further detailed as

$$\begin{aligned} \phi_0[n] &= \frac{1}{\sqrt{b\sqrt{\pi}}} e^{-n^2/2b^2} \\ \phi_1[n] &= -\frac{\sqrt{2}}{\sqrt{b\sqrt{\pi}}} \frac{n}{b} e^{-n^2/2b^2} \\ \text{and } \phi_2[n] &= \frac{1}{\sqrt{2b\sqrt{\pi}}} \left(\frac{2n^2}{b^2} - 1 \right) e^{-n^2/2b^2}. \end{aligned} \tag{5.12}$$

The parameters in (5.11) are determined as $a_0 = 1$, $a_1 = 2$, $a_2 = 2$, using least squares curve fitting of the recorded MIA signals during sports exercises. The parameter b of the MIA waveform depends on the type and intensity of the measured movement. It is defined as a uniform distribution between 0.1 and 0.5. In Figure 5.4, the variation in the value of b illustrates the alteration in the mean frequency of the MIA signal. The plot reveals that the spectral distribution of the added MIA is not confined to a specific frequency band and frequently overlaps with the sEMG spectral distribution.

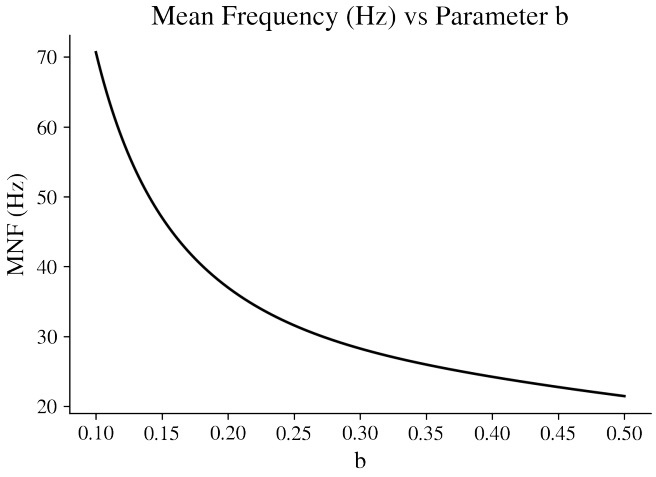


Fig. 5.4 The plot of the relationship between the mean frequency (Hz) of the MIA and the parameter b, as described in (5.12). The parameter b is uniformly selected from a range between 0.1 and 0.5, allowing for the exploration of the impact of varying levels of b on the mean frequency of the MIA.

5.7.2 Contamination of D_n with Synthetic MIA

In Section 5.6, the primary objective is to differentiate MIA windows without sEMG activity from those containing sEMG signals. Conversely, the focus of this section is to evaluate the model's ability to detect MIA contamination within active sEMG data streams. To achieve this, the sEMG windows in D_n, each 1024 samples in length, are artificially contaminated with MIA segments that are 128 samples long, positioned randomly within the windows. Contamination of sEMG data in D_n with synthetic MIA, compromising a maximum of 128 samples is given by

$$x[n] = \hat{x}[n] + \sum_{k=0}^{127} q[k] \cdot \delta[n - \tau - k], \tag{5.13}$$

where $x[n]$ denotes the sEMG signal with added MIA, $\hat{x}[n]$ represents the original, clean sEMG signal, $q[k]$ is the amplitude of the MIA at sample k, and $\delta[n-\tau-k]$ is the Dirac delta function, indicating the precise timing of the MIA. The variable τ is chosen randomly from the set $\{0,\ldots,896\}$, ensuring that the MIA is uniformly distributed across the sEMG window.

Given the adjustability of synthetic MIA amplitude, the amplitude of MIA is modulated so that the maximum MIA value aligns with the product of the adjustment rate and the average of the highest 5 rectified sEMG samples within the selected window. The amplitude adjustment of synthetic MIA utilised in (5.13), is calculated as

$$q[n] = \frac{\hat{q}[n]}{\max(\hat{q}[n])} \cdot AR \cdot \frac{1}{5} \sum_{i=0}^{4} \text{sort}_{\downarrow}(\hat{x}[n]), \tag{5.14}$$

where $\hat{q}[n]$ denotes the unnormalized MIA, and $\max(\hat{q}[n])$ is the maximum value of $\hat{q}[n]$ across all samples, ensuring normalization of the MIA. The term $\text{sort}_{\downarrow}(\cdot)$ refers to the sorting function that orders elements in descending order, applied to the sEMG signal $\hat{x}[n]$. The adjustment rate, denoted as AR,

is a predefined constant that scales the normalized MIA. This equalization of maximum MIA and maximum rectified sEMG amplitude complicates the differentiation of MIA within the sEMG stream. We apply 5 different adjustment rates and construct 5 different noisy datasets from D_n. It is worth noting that all the other variables τ, b are kept as the same over these 5 datasets for fair comparison. We use the signal-to-noise ratio (SNR) in (5.15) to compare the level of a desired signal $\hat{x}$ to the level of background noise q for various adjustment rates.

$$\text{SNR} = 10 \times \log_{10} \left(\frac{\sum_{n=1}^{N} \hat{x}[n]^2}{\sum_{n=1}^{N} q[n]^2} \right) \tag{5.15}$$

Owing to the synthetic MIA amplitude's tunability, it is modulated to ensure that its maximum value matches the product of the adjustment rate and the average of the top 5 rectified sEMG samples within the chosen window. This synchronisation of peak MIA and rectified sEMG amplitude heightens the challenge in differentiating MIA from the sEMG sequence. Therefore, detection precision is assessed over 5 distinct adjustment rates. The particular adjustment rates and their corresponding mean SNR values are presented in Table 5.5.

Table 5.5 Adjustment rates with corresponding mean SNR

Adjustment Rate	**0.8**	**0.9**	**1.0**	**1.1**	**1.2**
Mean (dB)	10.59	9.57	8.66	7.83	7.07

As expected, when we increase the magnitude of the noise with respect to the highest rectified sEMG, the noise becomes more dominant in the given sample. In the context of detecting the index of MIA within the sEMG window, a low SNR, which corresponds to a high *AR*, poses fewer challenges compared to a high SNR, i.e., a low *AR*. This is because, when the amplitude of the noise is lower than that of the sEMG data, the task of discriminating the noise from the signal becomes significantly more challenging.

In the context of the football dataset D_f, it is pertinent to note that no synthetic MIA is used. The dataset exclusively comprises natural MIA, which is inherently formed by real-life settings. Consequently, the SNR of the data within D_f is not quantifiable due to the absence of artificially added noise components. The identification and labeling of MIA indices within D_f have been meticulously performed through expert knowledge.

5.7.3 Experiments and Results

This study follows same preprocessing steps for both datasets. The preprocessing includes min-max normalization, DC-offset removal, and segmentation using a sliding window technique, with a window size of 1024. Both $\mathcal{D}_n$ and $\mathcal{D}_f$ datasets are configured to have the same sample size, enabling a fair comparison between the detection model's performance on synthetic and real-life MIA scenarios. This equivalence is achieved by adjusting the sample size of $\mathcal{D}_n$, where MIA is synthetically added, to

match that of $\mathcal{D}_f$. Each dataset encompasses a total of 12256 samples, divided into three subsets: 8576 for training, 2462 for testing, and 1214 for validation. This partitioning facilitates a thorough evaluation of the model's effectiveness in identifying MIA within sEMG signals. The experiments are conducted 10 times, each with a different random seed, and the results are subsequently averaged.

As illustrated in Table 5.6, the macro and micro F1 scores are presented for various adjustment rates within the synthetic dataset D_n. When the adjustment rate is 0.8, indicating that the MIA power is 20% below the maximum average of the sEMG signal, the macro F1 score achieves 93%. Lower adjustment rates pose challenges for MIA detection as the MIA power is less distinguishable from the sEMG signal. Our model proficiently identifies MIA zones with a macro F1 score exceeding 93%, even when the sEMG signal power surpasses the MIA. This demonstrates the efficacy of the BLSTM RNN model in distinguishing MIA zones.

To validate its performance on real-world data, the model is additionally trained and tested on D_f, consisting of data collected from athletes during training. The model attains a macro F1 score of 95%, substantiating its capability to accurately identify MIA zones in both synthetic and real-world scenarios during athletic activities.

Table 5.6 Comparison of macro and micro F1 scores at different adjustment rates for datasets D_n and D_f

Dataset	**Adjustment Rate**	**Macro F1 (%)**	**Micro F1 (%)**
D_n	0.8	93	98
D_n	0.9	94	99
D_n	1	95	99
D_n	1.1	96	99
D_n	1.2	96	99
D_f	-	95	99

5.8 Proposed U-Net Encoder Decoder (UED) Model for Denoising MIA from sEMG Signals

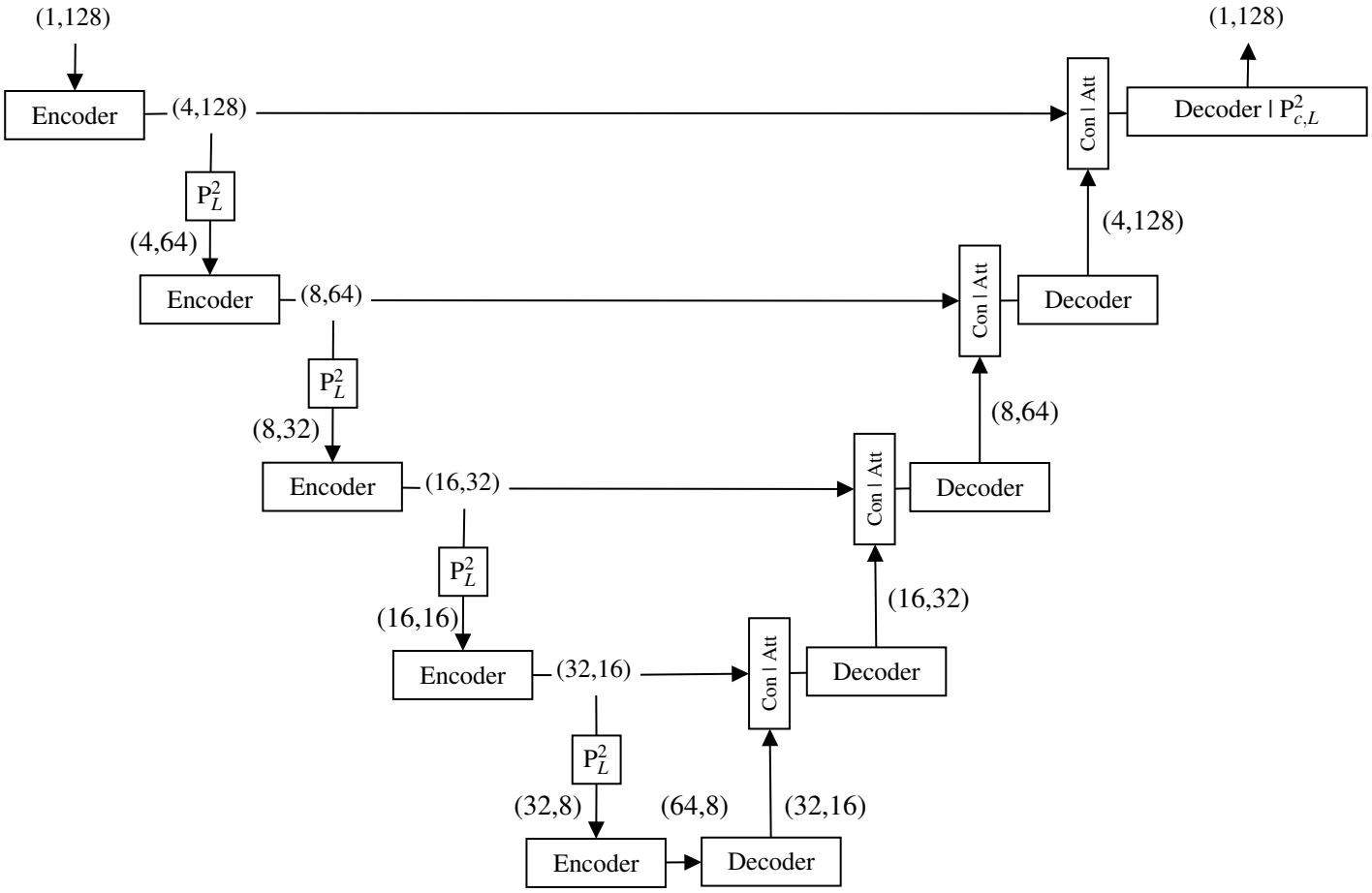


Fig. 5.5 The illustration of the block diagram for the network structure of the UED model. The model processes an sEMG window of $(1, 128)$, yielding a denoised vector of equivalent dimensions. Within the encoder block, the channel size is augmented by a factor of 4 in the first layer, and by a factor of 2 for subsequent layers, while PL^2 reduces the sequence length to half through length-wise average pooling. Post each Encoder, its output is fused to the corresponding level's decoder output via a copy and crop technique. The decoder block extends the length and reduces the channel size, while P$^2c, L$ performs average pooling by half for both channel and length dimensions. The Con | Att block denotes the concatenation over the channel axis followed by the application of the channel attention module.

The overall network structure for the U-net Encoder Decoder (UED) model is briefly illustrated in Figure 5.5. The model is designed to process an sEMG window of $(1, 128)$, subsequently producing a denoised vector of matching dimensions. The encoder block consists of an initial layer that quadruples the channel size, followed by subsequent layers that each double the channel size. Subsequent to each encoder block, the sequence length is halved via length-wise average pooling P_L^2. Following every Encoder sequence, the resultant output is concatenated via channel dimension with the decoder output of the corresponding level through a copy and crop methodology. The decoder block, conversely, increases the length and decreases the channel size. After the final decoder layer, average pooling is

performed using $P^2_{c,L}$, which halves both the channel and length dimensions. This reduction ensures that the output vector size matches the original input vector size. Lastly, the Con | Att block signifies the execution of concatenation over the channel axis, succeeded by the attention block as described in Chapter 4. For the UED model proposed in this section, the reduction ratio for the attention module is set to 4.

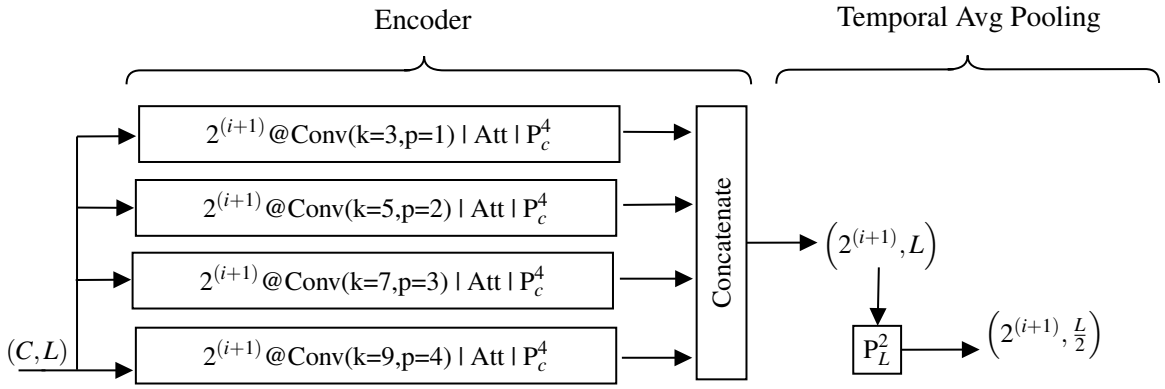


Fig. 5.6 The figure outlining the encoder block employed within the main model architecture, as referenced in Figure 5.5. The block receives an input of shape (C,L), representing channel size and sequence length, respectively, and yields an output tensor of shape $(2^{i+1}C,L)$, where i is the layer number. Four parallel convolution blocks are implemented in this model, each with an output filter size of 4 and varying kernel sizes—3, 5, 7, and 9. Here, 'Att' represents the channel attention block, and P^4_c denotes channel-wise average pooling that reduces the channel size by a factor of four. Post-encoder, the sequence length is halved via average pooling P^2_L. The model, by employing diverse kernel sizes and filter outputs, directs attention to each convolution block to discern the optimal channels.

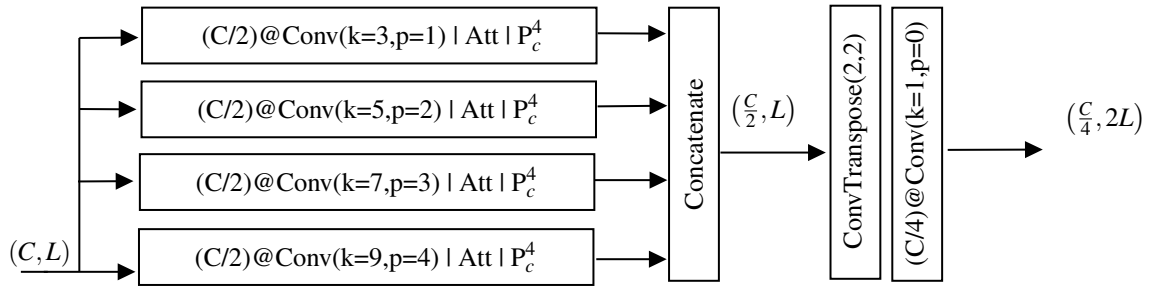


Fig. 5.7 The figure of a schematic representation of the decoder module, referenced in Figure 5.5. The module receives an input tensor of dimensions (C,L) and outputs a tensor of dimensions $(\frac{C}{4},2L)$. Essentially, the module performs operations inverse to those of the encoder, modifying the shape of the vector correspondingly. Mirroring the actions of the encoder block, this model deploys 4 parallel convolution blocks, succeeded by channel attention and P^4_c for channel-wise average pooling to size $\frac{C}{8}$. A distinction is that the convolution blocks employ an output filter size that is half of the input channel size. The concatenated vector is then integrated with the convolution transpose block, utilizing a stride and kernel sizes of two, thus increasing the sequence length by 2. The final convolution block further reduces the channel size to $\frac{C}{4}$.

The encoder block, depicted in Figure 5.6, utilizes four parallel convolution blocks, each with kernel sizes of 3, 5, 7, and 9. The output filter size for each block is determined by the expression 2^{i+1}, where i is the layer number. For the first layer, the channel size is increased by a factor of four,

and for subsequent layers, the channel size is doubled. The integration of various kernel sizes aims to discern distinct patterns across different frequency bands, facilitating precise localization and extraction of MIA. A larger kernel size captures low-frequency alterations, while a smaller one distinguishes high-frequency transitions. Each output undergoes a fusion to the channel attention module after the convolution block, followed by a degree 4 channel-wise average pooling. This represents a weighted average for each convolution block output, with weights allocated through an adaptive attention module, pinpointing the most impactful channels. The pooled outputs are then concatenated over the channel dimension, resulting in an output tensor of shape $(2^{i+1}C, L)$ after the encoder, where the sequence length is subsequently halved via average pooling P_L^2.

Contrary to the encoder, the decoder block extends the sequence length and reduces the channel size. As illustrated in Figure 5.7, the module initiates with 4 parallel convolution blocks, the output filter size of which is equal to half of the input channel size. Similar to the encoder block, these convolution blocks are succeeded by a channel attention module and channel-wise average pooling. Once the outputs of the convolution blocks are concatenated over the channel dimension, a 1-dimensional convolutional transpose block with a stride and kernel size of 2 is employed. This operation effectively doubles the sequence length, functioning as an inverse convolution operation with a stride of 2. Post sequence length upsampling, a convolution block with a kernel size of 1 is applied, yielding an output channel size of $\frac{C}{4}$, which results in an output vector shape of $(\frac{C}{4}, 2L)$.

5.9 Validation Experiments for the Denoising Framework

In this experiment, the efficacy of the denoising model is evaluated using the D_n dataset, enriched with synthetic MIA. As shown in Section 5.7, synthetic MIA is employed for testing and comparing the denoising efficacy of the proposed methodologies, since having knowledge of both clean and noisy signals is compulsory for calculating SNR improvement.

A comparative analysis is conducted between the proposed models and the prevailing state-of-the-art models, utilizing the output from the detection model. Contemporary techniques employed to evaluate and benchmark the UED model performance are outlined, as the references for these models are provided. These models have been adapted to suit the sEMG data used in this study. The identified MIA region is expanded to encompass 128 samples, aligning with the prerequisite of all tested models to include a noisy signal comprised of 128 samples. The UED model is also tested with various loss functions: Mean Square Error (MSE) and Soft-DTW. MSE outperforms Soft-DTW with 1.3 dB SNR improvement, to this end, MSE is selected as the loss function for all the models tested.

The effectiveness of the denoising method, UED, is assessed through four performance metrics: MSE, Mean Absolute Error (MAE), Signal-to-Noise Ratio Improvement (SNR_{imp}), and Percent Residual Difference (PRD), where $\hat{x}$ is the clean signal, $\bar{x}$ is the denoised signal and x is the noisy sEMG stream.

SNR_{imp} quantifies the enhancement in SNR between the noisy signal and the denoised signal, represented as:

$$\text{SNR}_{\text{imp}} = \text{SNR}_{\text{after}} - \text{SNR}_{\text{before}}, \tag{5.16}$$

where $\text{SNR}_{\text{after}}$ is the SNR of the denoised signal, i.e., model output, and $\text{SNR}_{\text{before}}$ is the SNR of the noisy signal shown as

$$\begin{aligned} \text{SNR}_{\text{before}} &= 10 \times \log_{10}\left(\frac{\sum_{i=1}^{N} \hat{x}_i^2}{\sum_{i=1}^{N} (\hat{x}_i - x_i)^2}\right) \\ \text{SNR}_{\text{after}} &= 10 \times \log_{10}\left(\frac{\sum_{i=1}^{N} \hat{x}_i^2}{\sum_{i=1}^{N} (\hat{x}_i - \bar{x}_i)^2}\right) \end{aligned} \tag{5.17}$$

PRD serves as a normative measure for the quality of the reconstructed signal, with lower values indicating improved quality:

$$\text{PRD} = \frac{\sqrt{\sum_{i=1}^{N} (\hat{x}_i - \bar{x}_i)^2}}{\sqrt{\sum_{i=1}^{N} \hat{x}_i^2}} \times 100 \tag{5.18}$$

The experimentation is extended to investigate whether segmenting the problem into detection and denoising subproblems enhances performance or if a unified model—processing an input without predefined knowledge of the MIA regions and denoising the signal—is more successful. To substantiate this, the proposed model is deployed to process an input comprising 1024 samples, absent of any predefined knowledge of MIA zones, and exclusively the MIA zones.

5.10 Results

In Table 5.7, the denoising efficiency of the proposed model is presented, demonstrating variations across different depths: 3, 4, and 5. The UED model, at a depth of 5, displays the best denoising performance. Moreover, the integration of channel attention elevates the SNR by 0.26 dB. Consequently, a structure consisting of 5 layers with attention is designated as our primary model, as depicted in Figure 5.5.

Table 5.7 Comparative analysis of performance metrics across varied depth sizes and attention mechanism for the proposed UED model

Model	**MAE** (10^{-4})	**MSE** (10^{-4})	$\textbf{SNR}_{\textbf{imp}}$
3 Depth	173.2	15.7	14.05
4 Depth	148.4	8.3	15.15
5 Depth	119.8	7	16.94
5 Depth w/ Att.	**112.6**	**6.6**	**17.20**

In Table 5.8, a comparative analysis between the proposed UED model and prevailing denoising models is presented. This study introduces a framework where denoising occurs after detection by the BLSTM RNN model. For a balanced comparison, both the state-of-the-art models are evaluated with and without detection through BLSTM RNN. The combined approach of BLSTM RNN and UED showcases a notable improvement, achieving a 17.20 dB SNR_{imp}, surpassing the top-performing state-of-the-art model, CPDAE, by 7.01 dB SNR_{imp}.

To ensure an accurate and fair evaluation of the UED model's performance, the BLSTM RNN detection framework (referred to as the 'on' case in Table 5.8) is applied to all models under comparison for denoising performance. The UED model distinctly surpasses all other models, achieving a 17.20 dB SNR_{imp} and 18.50 PRD. This marks a 1.53 dB higher SNR_{imp} than the next best model, CPDAE, as cited in [90]. Additionally, it elevates the average SNR_{imp} of the state-of-the-art models by 9.69 dB.

Table 5.8 Comparative analysis of performance metrics for contemporary methods and the proposed UED model with and without BLSTM RNN detection

Model	**Framework**	**MSE** (10^{-4})	**SNR_{imp}**	**PRD**
Filter [39]	off	-	0.44	-
Filter [39]	on	110.2	1.92	34.07
FFT-based [100]	on	69.2	4.95	24.62
CPDAE [90]	off	-	10.19	-
CPDAE [90]	on	8.6	15.67	21.25
UED	off	-	14.08	-
UED	on	**6.6**	**17.20**	**18.50**

To determine if the use of a detection model before a denoising model enhances the denoising performance, all models are tested without the BLSTM RNN detection model. This approach lets the models directly accept an input of 1024 samples without pinpointing the MIA region. The denoising SNR witnesses improvements of 3.12 dB, 5.48 dB, and 1.48 dB for UED, CPDAE, and Filter, respectively. This emphasises that the proposed method of problem segmentation, combined with distinct models for detection and denoising, significantly boosts performance.

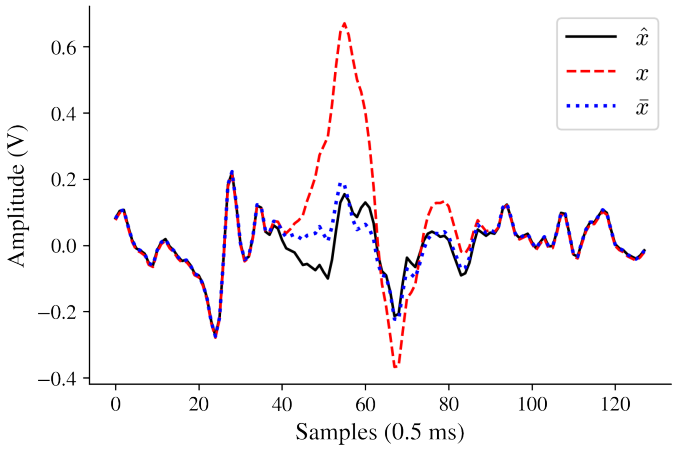


Fig. 5.8 This plot shows the original noisy signal, x, the clean signal with no MIA, $\hat{x}$, and the denoised signal from the UED model, $\bar{x}$. It provides a clear comparison, illustrating how the UED model improves signal SNR by denoising the recorded signal.

Figure 5.8 depicts the noisy signal, x, the noise-free signal, $\hat{x}$, and the denoised signal processed by the UED model, $\bar{x}$. This representation demonstrates the UED model's capability to enhance signal SNR by effective denoising of the recorded signal. The reconstructed signal closely mirrors the clean signal, signifying the model's efficacy in eliminating MIA while preserving the essential sEMG data.

5.11 Conclusion

This research introduces a comprehensive framework designed to denoise MIA within sEMG streams and contrasts its effectiveness with existing state-of-the-art denoising models across synthetic and real-world datasets.

The findings of this study highlight that for the denoising of MIA in sEMG, employing two distinct models—one to identify the MIA region within the sEMG stream and the other to denoise the detected region—is more advantageous than leveraging a singular model. The separation of the task into two specific subproblems enhances the denoising performance by 3.12 dB SNR_{imp}.

Moreover, this research validates that the BLSTM RNN, mainly used for motion classification and onset detection in the sEMG research domain, is also useful for noise detection. By applying both synthetic and real-world datasets, the study not only validates the performance of the noise detection model but also corroborates the MIA mathematical model proposed in [55]. The model showcases a high F1 score and precision, especially when the MIA power is similar to the sEMG signal. This shows the model's capability to pinpoint MIA zones even in challenging scenarios where noise power is less than the sEMG. A review of the implications of model depth and the attention mechanism, outlined in Table 5.7, discloses that a configuration comprising 5 layers with attention is optimal, improving SNR by 0.26 dB. Furthermore, the comparative analysis presented in Table 5.8 underscores the unmatched denoising proficiency of the UED model against its contemporaries, exhibiting remarkable improvements in SNR_{imp} and PRD.

The existing MIA model exhibits two key limitations: first, it relies solely on a single scaling parameter b in (5.11) as the source of randomness. This simplification could potentially make it easier for deep learning models to detect and denoise MIA, as they only need to adapt to a single variability factor. Second, the model assumes that noise is purely additive, which fails to account for multiplicative noise effects that can occur, for example, from shifts in electrode positions relative to the EMG source. Such multiplicative noise can significantly alter the signal's integrity in ways not captured by an additive noise model. These limitations can potentially be addressed through the incorporation of a Generative Adversarial Network (GAN). In this setup, the discriminator is trained to differentiate between sEMG with real MIA contamination and sEMG with synthetic MIA contamination, while the generator is designed to contaminate the sEMG windows with MIA realistically. Such an approach could mitigate the MIA model's reliance on a single parameter for randomness and address limitations arising from the assumption that noise is purely additive, without considering multiplicative noise effects like electrode position shifts.

Chapter 6

Motion Classification

This chapter delineates the motion classification module of the proposed framework, segmented into two distinct parts. The initial segment unveils the fundamental deep learning model crafted for motion classification. Additionally, the performance of this model is benchmarked against leading motion classification models using the well-established D_n dataset. The model is also tested on our D_f dataset to evaluate its applicability and robustness in real-life sports settings. The latter segment refines the aforementioned model, tailoring it to better fit sports-related motion classification tasks when working with limited datasets.

6.1 Deep Learning Model for Motion Classification

6.1.1 Introduction

Motion classification via sEMG refers to the process of identifying and categorising distinct muscle movements by analysing sEMG signals. Motion classification techniques can also be utilised to distinguish sports-related drills using only the sEMG signals. Muscle dynamics vary with different movements, making motion classification using sEMG signals indispensable in sports settings. Most studies of sEMG-based motion classification utilise the D_n dataset for hand gesture recognition (HGR) [88]. This chapter extends the focus beyond the D_n dataset to include the D_f dataset, emphasising sports movements involving the hamstring, adductor, and soleus muscles. The aim is to enhance sEMG-based HASI evaluation by classifying movements targeted at these muscles. Testing the model's performance on the D_f dataset, which features real-life sports settings, offers a significant advancement in risk assessment and injury prevention.

Rather than processing the full raw data set, the model decomposes the input signal into two distinct segments, each corresponding to specific frequency bandwidths: the high (100 Hz to 300 Hz) and low (0 to 100 Hz) frequency bands. Signals from both these bands are introduced into the model in parallel. Subsequently, an attention block integrates the combined output. This approach facilitates the model in selectively extracting features inherent to each frequency band and then focusing on critical attributes from these extracted features. Such a mechanism enables the model to comprehend nuances across diverse frequency regions, accentuating key features in the process. This strategy leverages

the physiological differences in muscle activation patterns captured within these frequency bands, thereby improving the model's ability to generalize across varied datasets. This methodology is not based on any prior work in the literature; it was created solely with the motivation of enhancing model performance and adaptability by focusing on distinct frequency bands. To address the variability in kernel and depth sizes within the feature extractor, a refined version of the KD block, as introduced in Chapter 4, is utilised.

6.1.2 State-of-the-art

The classification of motion using sEMG signals can be approached through two principal ways: feature engineering and feature learning. Each procedure utilises feature vector representations of sEMG signals. Feature engineering is based on signal processing algorithms, whereas feature learning involves deep neural networks [88, 175, 151]. Studies in hand gesture recognition indicate a higher efficacy of feature learning over feature engineering in classifying sEMG-based motions [8, 93, 58, 136].

For sEMG data, feature learning predominantly employs CNNs to extract distinct feature representations [175]. Research demonstrates that CNNs excel in classifying various hand gestures, surpassing the performance of feature engineering techniques using support vector machines [130]. Typically, sEMG raw data is converted into a matrix format, which is then processed by two-dimensional (2D) CNNs [187, 8, 67]. Another approach is to use sEMG spectrograms, created using the short-time Fourier transform (STFT), as inputs for 2D CNNs, capturing both time and frequency domain information [3, 46]. A comparative analysis shows that the continuous wavelet transform (CWT) is more effective than STFT when used with CNNs [34]. The concept of slow-fusion CNNs is also introduced, enhancing feature learning by establishing temporal correlations between CWT images.

Further advancements are seen in a hybrid CNN-RNN architecture introduced by Hu et al., where sEMG data is segmented temporally and processed through parallel 2D CNNs before being fed into an RNN with an attention mechanism [82]. The EMGHandNet model integrates a 1D CNN with a BiLSTM-RNN, showing improvements over Hu et al.'s approach and emphasising the role of 1D CNN in sEMG data processing [94]. Rahimian et al. advance this further with a dilated 1D CNN block, achieving better results in certain datasets [137]. Recently, the TraHGR-Huge model using a transformer-encoder structure with multi-head attention has been developed. It employs two parallel transformer structures to process normal and transposed inputs, facilitating the learning of channel-wise relationships and marking significant progress in HGR [183].

6.1.3 Contributions

The primary contributions of this section are outlined as follows:

- A motion classification model is applied to a dataset collected from professional football players, presenting a unique contribution to sEMG-based sports science evaluations.
- The proposed model surpasses the state-of-the-art by achieving a 3.29% accuracy improvement on the D_n dataset and an 8.66% improvement on the $D_{n,b}$ dataset.

- The presented model, which outperforms all others in HGR, is also tested on a D_f dataset encompassing lower extremity muscles, professional athletes, and a different sensor type. The accuracy reaches to 96.8% without any parameter modifications.
- The adoption of the 1D CBAM, in place of the conventional MHA for sEMG applications, is utilised to highlight essential features in both channel and temporal dimensions.
- Frequency-split feature extraction technique is introduced for the first time in sEMG motion classification. This strategy processes sEMG data through parallel models across diverse frequency bands, enabling the model to selectively focus on frequency-specific attributes.
- The model proposed is universally applicable for all sEMG-based motion classification tasks. It utilises five kernel sizes, three depth sizes, and an attention mechanism. Such a configuration amplifies the model's versatility, making it suitable for a range of sEMG-based tasks without necessitating alterations to the kernel or hidden layer sizes.

6.1.4 Proposed Model

Improved outcomes in motion classification models are often observed with preprocessed sEMG data, as documented in [67, 8, 167]. This study adopts preprocessing steps from [136], encompassing Butterworth filtering, normalization, and segmentation. To enhance the model's classification accuracy, an attention mechanism is applied that targets both the high and low-frequency sEMG data bands. This requires two separate preprocessing stages to isolate high-frequency band sEMG, marked as x_h, and low-frequency band, x_l. Details of the preprocessing steps, alongside the sizes of the training and test sets for each dataset ($\mathcal{D}_n$, $\mathcal{D}_{n,b}$, and $\mathcal{D}_f$), are presented in Table 6.1.

The strategy for frequency band separation in this study focuses on standardising the power spectral density across sampled signals from the dataset D_f. Specifically, the low-frequency band encompasses frequencies from 0 to 100 Hz, while the high-frequency band spans 100 to 300 Hz. This categorisation is based on observations that frequencies within the $300-500$ Hz range contribute only a minimal portion (approximately 0.02) of the total power in sEMG samples. Furthermore, the median frequency of power spectral density in our dataset is around 96 Hz, consistent with research indicating that the median sEMG frequency typically falls between 80 and 120 Hz [105, 106]. By equalising power spectral density, the twin network architecture of our model processes each frequency band effectively, ensuring a balanced complexity in signal analysis for more efficient feature extraction. Segmenting the signal into these defined frequency bands improves the system's ability to analyse data and enhances the extraction of distinct signal characteristics.

Preprocessing sEMG with high-frequency band x_h

The signal is processed through a 9th order band-pass Butterworth filter, set between 100 and 300Hz as cutoff frequencies. Following this, the filtered signal undergoes a transformation using the μ-law

algorithm, as

$$x_\mu = \text{sign}(x)\frac{\ln(1+\mu|x|)}{\ln(1+\mu)}, \tag{6.1}$$

where $\mu = 2048$. Subsequent to the μ-law transformation, minmax normalisation is applied to scale the data given as

$$x_{\text{norm}} = \frac{x_\mu - x_{\mu,\min}}{x_{\mu,\max} - x_{\mu,\min}}. \tag{6.2}$$

The normalized signal is then segmented through the sliding window technique, with the window size and overlap parameters as detailed in Table 6.1.

Preprocessing sEMG with low-frequency band x_l

To eliminate high-frequency noise above 100Hz, a 9th-order Butterworth low-pass filter is applied. For further refinement of the data, a Savitzky-Golay filter is employed. This filter is defined by a window size of $W_n = 51$ and a polynomial order of $P_o = 3$, details of which are elaborated in [54]. This filter adeptly smooths the data while preserving essential signal characteristics, as noted in [146]. After the application of these filters, the normalisation and segmentation steps are applied as outlined in Section 6.1.4.

Table 6.1 Comprehensive summary of preprocessing steps and training/test set sizes for datasets D_n, $D_{n,b}$, and D_f

	D_n	$D_{n,b}$	D_f
Window Size	400	400	200
Overlap	0	200	0
PLN filter	yes	yes	no
Wavelet denoising	yes	yes	yes
Butterworth filter	yes	yes	yes
u-law	yes	yes	yes
min-max	yes	yes	yes
Training set size	192393	121168	198048
Test set size	95923	60259	23224

Network Architecture

In the proposed architecture, depicted in Figure 6.1, the model processes three-dimensional inputs x_h and x_l. These inputs are characterised by their batch size, channel number, and window length dimensions. Each input signal is independently fed into the KD block, which is the feature extractor. The outputs from these KD blocks are then amalgamated along the channel axis using a concatenation block, leading to a combined output in the channel dimension. Subsequently, this merged output undergoes processing by a 1D convolution block, which utilises a kernel and stride both set to 1.

An attention block enhances the model's focus on pivotal channels and samples. After the attention block, the model embarks on a series of adaptive average pooling layers and convolution blocks. This sequence is strategically designed to reduce the signal's channel size and sequence length, aligning them with the designated number of prediction classes. The employment of distinct KD blocks for processing the high and low-frequency bands of the signal is a calculated choice. It aims to direct the model's attention towards varying frequency bandwidths. After undergoing channel-wise concatenation and the 1D convolution process, the attention block is responsible for learning and focusing on the specific features extracted from these distinct frequency bands.

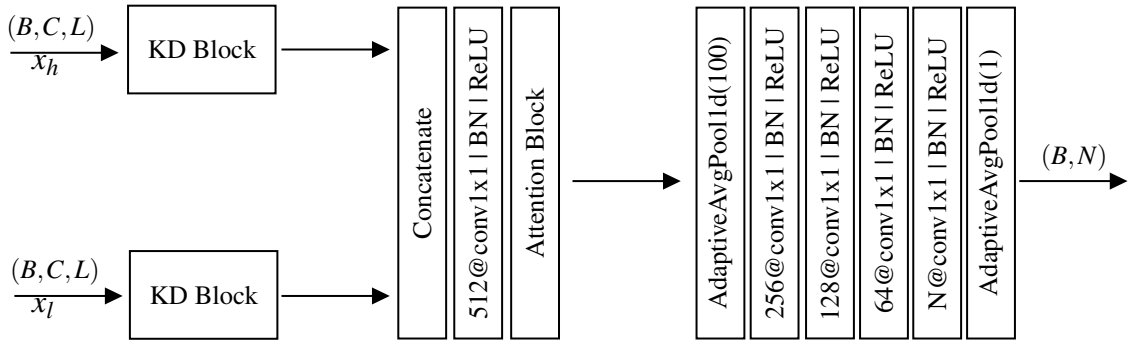


Fig. 6.1 The schematic representation of the proposed model. The model accepts two input tensors of shape (B,C,L), corresponding to low frequency-band, x_l, and high frequency-band, x_h. B corresponds to the batch size, C refers to the number of channels, and L indicates the sequence length. Convolution layers are denoted as $\mathbf{f}@Conv\mathbf{k} \times \mathbf{s}$, where f, k, and s represent the output filter size, kernel size, and stride, respectively. The variable N indicates the number of classes in the dataset, with specific values for different datasets: 49 for D_n, 17 for $D_{n,b}$, and 9 for D_f.

The KD block, D block, and attention modules are introduced in Chapter 4. Reference to Figure 6.2 highlights the modifications made to the KD block for the model presented in this chapter. Specifically, the stride is adjusted to 2 to reduce the signal length, and the initial 1D convolution has an output channel size of 16. Additionally, a 1D max pooling with a kernel size of 3 and a stride of 1 is incorporated prior to a 1D convolution block with a stride set to 2. This is done to align with the output shape of the D Block. Consequently, the KD block delivers a channel output of 256, in contrast to the 512 found in Chapter 4. Within the D block, only the output filter size is altered, set to 16, as compared to the 32 in Chapter 4. The reduction ratio is selected as 16 in the KD Block and 128 in the final attention module as shown in Figure 6.1. Throughout the training process, the model employs the Binary Cross-Entropy (BCE) loss and utilises the ADAM optimiser.

6.1.5 Experiments & Results

In this section, experiments are conducted on three datasets: D_n, $D_{n,b}$, and D_f, with accuracy serving as the evaluation metric. Each dataset is divided into training, validation, and test subsets, with specifics detailed in Table 6.1. The training subset informs the model's learning, the validation subset aids in avoiding overfitting, and the test subset facilitates method comparisons. These experiments are executed ten times using varied random seeds, and the outcomes are subsequently averaged. The

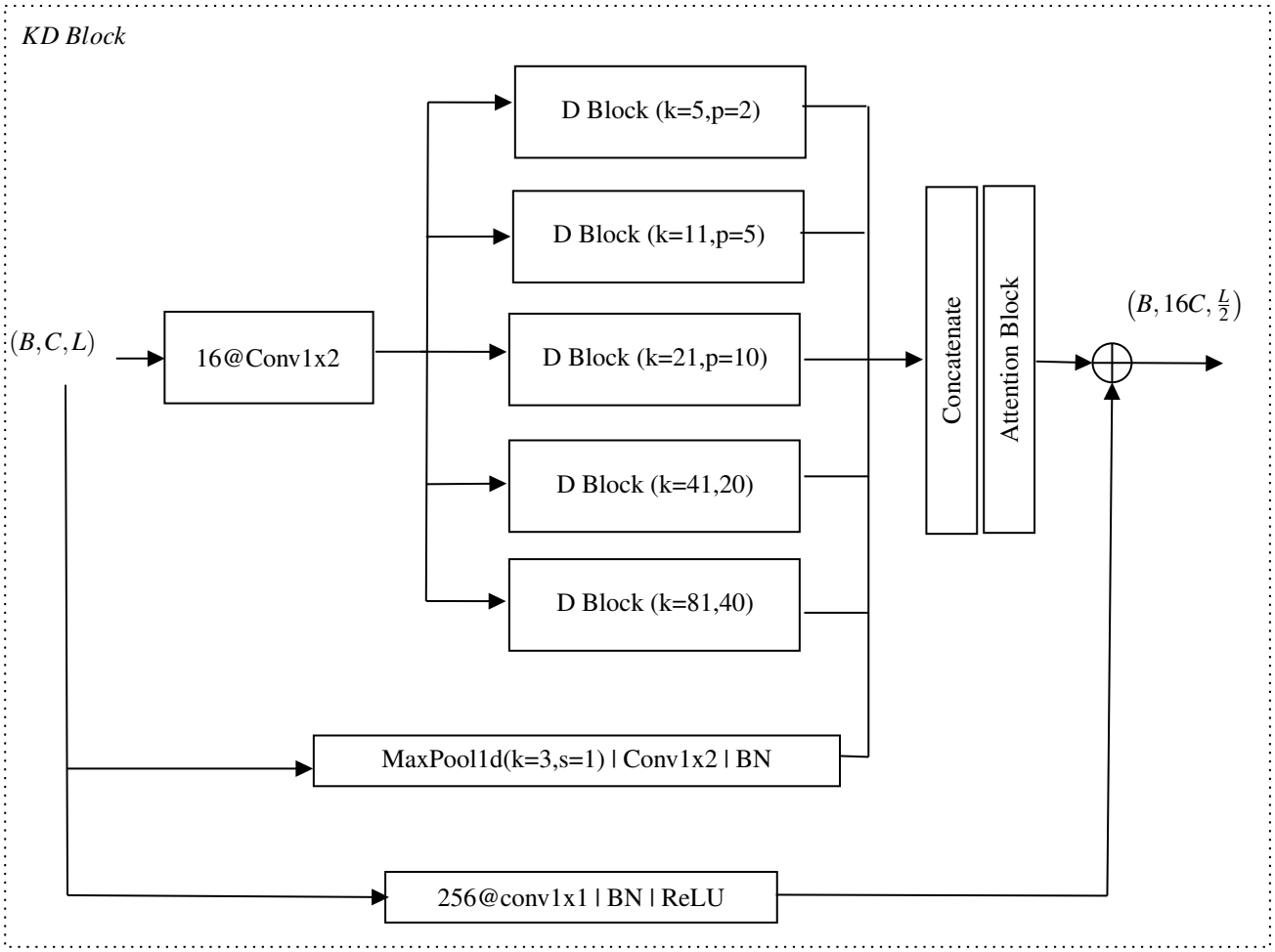


Fig. 6.2 The schematic representations of the KD block. The KD block, presented in Chapter 4, has undergone slight modifications: notably, the stride is set to 2 to downsample the signal length, and an initial 1D convolution has an output channel size of 16. As a result, the KD block's channel output is 256, compared to the original 512. The block integrates an initial 1D convolution, parallel D blocks with diverse kernel sizes, a concatenation process with a residual tensor, an attention segment, and a summation with another residual. This transforms the input tensor from a shape of (B,C,L) to $(B,16C,\frac{L}{2})$.

proposed model is benchmarked against state-of-the-art models on the datasets D_n and $D_{n,b}$. Accuracy scores of each model are summarized in Table 6.2. To ensure a fair comparison, identical training and testing repetitions and a consistent window size of 200 ms are used. Following standard practice, the model is trained using repetitions 1, 3, 4, and 6 and tested using repetitions 2 and 5, as outlined in Chapter 3. No parameter tuning is conducted on the model for all tests across various datasets. The only adjustment made is to the final output layer size, aligning it with the specific number of classes in each dataset: 49 for D_n, 17 for $D_{n,b}$, and 9 for D_f.

In the D_n dataset, the developed model achieves a classification accuracy of 95.30%, surpassing its closest contemporary competitor by 3.29%. When benchmarked against seven contemporary studies, the model's accuracy significantly exceeds their mean by 15.7% and median by 13.1%, representing a major advancement. Within the $D_{n,b}$ dataset, our model records an impressive 98.80% accuracy,

Table 6.2 Comparison of the proposed model's accuracy with other state-of-the-art models across the D_n, $D_{n,b}$, and D_f datasets, evaluated at a 200 ms window size.

Model	Dataset	Accuracy (%)
CNN by Aztori et al. [8]	D_n	60.30
Feature engineering [9]	D_n	75.30
CNNLM [26]	D_n	78.71
CNN-RNN by Hu et al. [82]	D_n	82.20
CNN by Wei et al. [166]	D_n	83.70
TraHGR-Huge [183]	D_n	86.18
EMGHandNet [94]	D_n	92.01
The proposed model	D_n	**95.30**
SVM by Zhai et al. [185]	$D_{n,b}$	81.07
CNN by Zhai et al. [185]	$D_{n,b}$	82.22
CNN by Ding et al. [44]	$D_{n,b}$	83.79
TraHGR-Huge [183]	$D_{n,b}$	88.91
CNN by Rahimian et al. [137]	$D_{n,b}$	90.14
The proposed model	$D_{n,b}$	**98.80**
The proposed model	D_f	**96.30**

setting a new standard in the field. This result not only outperforms the nearest competing CNN model by Rahimian et al. [137] by 8.66% but also surpasses the average and median accuracies of similar models by 13.57% and 15.01%, respectively.

For the D_f dataset, the model attains a 96.30% accuracy rate in distinguishing 9 distinct hamstring exercises. Detailed classification metrics such as precision, recall, and F1-score for these 9 exercise categories are tabulated in Table 6.3, corresponding to the exercise descriptions in Section 3.2. Notably, the model attains ideal F1 scores for exercises 2 through 5, attributed to their primarily isometric nature with minimal isotonic contractions. Conversely, exercises 6, 7, and 9, characterized by more dynamic isotonic movements, achieve lower F1 scores, highlighting the challenges in discerning more complex motion patterns within the chosen 200 ms window.

The model's classification efficacy is visually depicted through confusion matrices for both $D_{n,b}$ and D_f, as shown in Figures 6.4 and 6.3, respectively. These matrices offer an in-depth perspective of the model's accuracy across various classes. Figure 6.4 and Table 6.4 display the model's performance across the 17 classes of $D_{n,b}$. In Figure 6.4, the diagonal elements, indicative of accurate class predictions, highlight the model's precision. The lower off-diagonal elements, which represent misclassifications, confirm its ability to reliably categorize instances across diverse classes. The colour gradient in the matrix visually emphasizes the proportion of correct and incorrect classifications, further illustrating the model's consistent classification accuracy. Similarly, Figure 6.3 shows the confusion matrix for $\mathcal{D}_f$, comprising 9 classes. The confusion matrix for D_f also exhibits a significant number of

correct predictions, evident in the diagonal elements. The consistency in correct classifications across various classes suggests the model's unbiased performance, ensuring high predictability for each class.

Table 6.3 Summary of classification metrics for a multiclass prediction model, displaying precision, recall, and F1 score for 9 classes in D_f.

Class Name	Precision	Recall	F1 Score	Support
1. 60° Adductor Squeeze	0.95	0.97	0.96	2290
2. Hamstring Claw	1.00	1.00	1.00	2358
3. Hip Extension	0.99	1.00	1.00	2323
4. Hip Flexion	1.00	1.00	1.00	2288
5. Sitting Soleus Raise	1.00	1.00	1.00	2241
6. SL Bent-Knee Soleus Raise	0.96	0.95	0.95	2822
7. SL Elevated Glute Bridge	0.94	0.92	0.93	3376
8. Isometric Prone Squeeze at 0°	0.97	0.98	0.97	2417
9. SL Prone Curls	0.93	0.93	0.93	3109

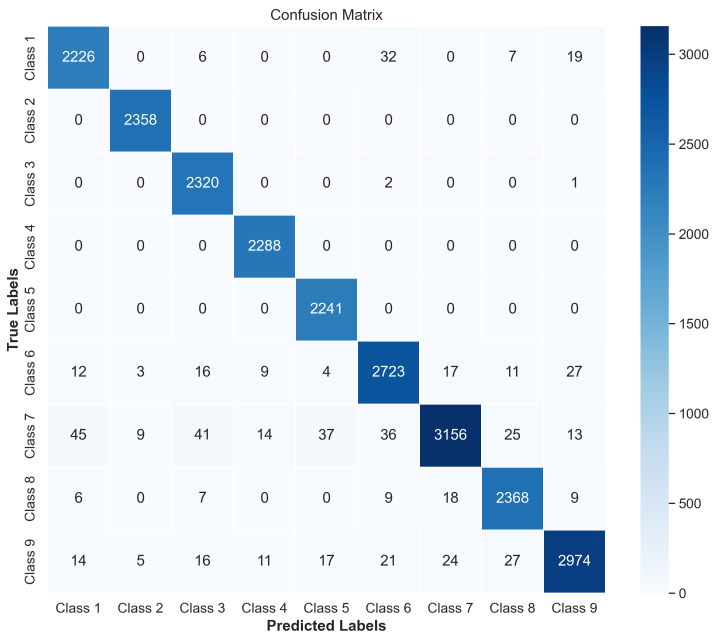


Fig. 6.3 Confusion matrix illustrating the classification performance for 9 classes in the D_f dataset. Correct predictions for each class are indicated by diagonal elements, whereas misclassifications are represented by off-diagonal elements. The color gradient in the matrix highlights the number of instances in each category, effectively demonstrating the model's accuracy and discriminatory capability in classification tasks.

Table 6.4 Summary of classification metrics for a multiclass prediction model, displaying precision, recall, and F1 score for 17 classes in $\mathcal{D}_{n,b}$.

	precision	recall	F1 score	support
1	1.00	1.00	1.00	4722
2	1.00	0.91	0.95	4070
3	0.99	0.99	0.99	3691
4	0.99	1.00	0.99	3522
5	0.99	1.00	0.99	3639
6	0.96	1.00	0.98	2967
7	0.96	1.00	0.98	4519
8	0.99	1.00	1.00	3431
9	1.00	0.99	0.99	3701
10	0.99	0.99	0.99	3809
11	0.99	1.00	0.99	4141
12	0.99	0.99	0.99	3353
13	0.99	1.00	0.99	2689
14	0.99	0.99	0.99	2713
15	1.00	0.99	0.99	3285
16	0.98	0.99	0.99	3292
17	0.98	0.99	0.98	2715

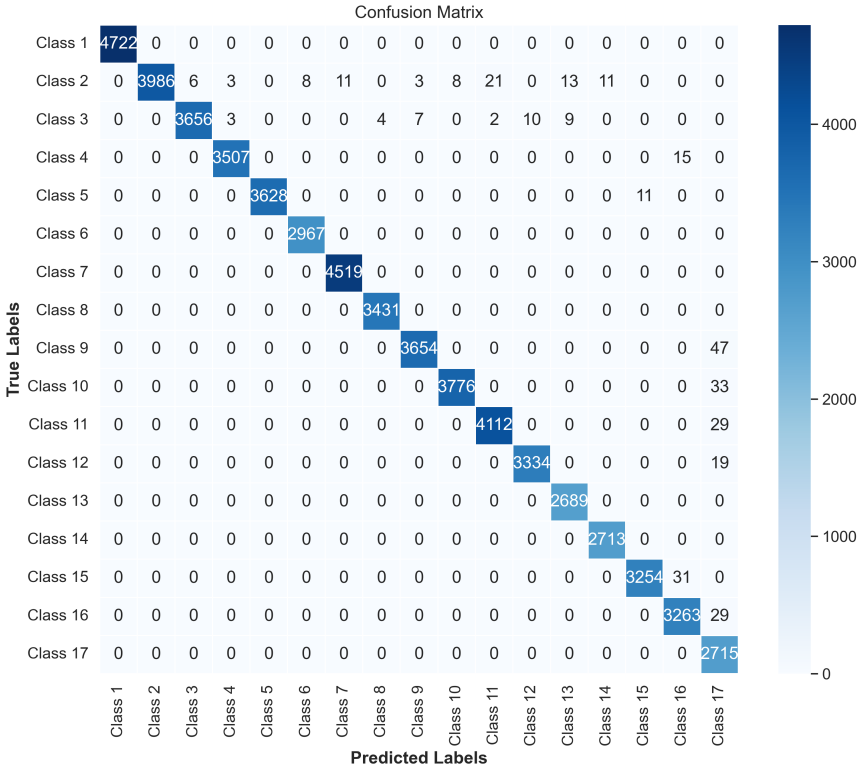


Fig. 6.4 Confusion matrix depicting the classification accuracy over 17 classes in the $D_{n,b}$ dataset. Diagonal elements indicate the correct predictions per class, whereas off-diagonal elements represent misclassifications. The color gradient visually demonstrates the volume of instances in each category, underscoring the model's accuracy and precision in distinguishing between classes.

6.2 Adapting Deep Learning for Constrained Datasets

6.2.1 Introduction

In the preceding section, the proposed model achieves an accuracy of 96.30% on a sports dataset. This performance requires a large training dataset that includes a significant number of equivalent independent observations. Specifically, these observations are derived from various subjects across different sessions, a critical consideration in the sEMG motion classification domain. The deep learning models, when trained using conventional approaches, require a vast and significant number of equivalent independent observations because they rely on generalization over a diverse dataset. This requirement entails capturing a wide range of subject and session variations, as differences in subjects and sessions are substantial in domains like sEMG motion classification. However, acquiring such extensive datasets with high-quality independent observations is often challenging in real-time sports applications.

Collecting sufficient data in sports settings is challenging for several reasons. First, there are hundreds of different exercises, i.e., motion classes, and obtaining a large dataset for each potential class is not feasible [40]. Secondly, even when focusing on a few essential exercises, collecting adequate data from elite athletes is constrained. Their tight schedules often do not permit extensive data collection.

This section introduces few-shot metric learning (FSL) techniques to address the dataset size constraints in sports scenarios. FSL, inherently designed for minimal data, aligns with the challenges in sports, where amassing vast datasets is problematic [165]. Employing metric learning endows the model with the ability to learn a similarity function, accepting pairwise inputs to discern the similarity or difference between data points [95]. This strategy, reminiscent of methods observed in Siamese and Relation networks and concepts like triplet loss, enables the model to discern comparative features [101, 152, 160, 155].

In the competitive domain of sports, teams often guard their data, rendering the prospect of a publicly accessible sports database similar to Ninapro quite unlikely [40]. Recognising this challenge, this study adopts an alternative method. Instead of depending solely on sports-specific datasets during model training, the proposed methodology leverages the publicly available HGR dataset, Ninapro. Using this dataset, the model is trained via the FSL techniques to learn a similarity function. FSL usually utilises a "support set", a small group of examples for learning, and a "query" for prediction [165, 155]. Building upon this concept, the model is meticulously fine-tuned with a limited number of soccer data samples — 5, 10, 20, and 50 shots. After this fine-tuning, the model's performance is evaluated using the comprehensive soccer dataset. Notably, this approach enables achieving significant accuracy on the extensive soccer dataset with just a few sample data points.

6.2.2 State-of-the-art

Several recent studies employ domain adaptation and few-shot learning (FSL) to address the challenges caused by sEMG data variability. This variability often arises from electrode shifts [154, 181, 182],

muscle fatigue [68, 115], sweat [1], varying sensor placements [91, 58, 23], and differences in individual physiology [98]. On top of that, Hug et al. suggest that muscle recruitment strategies not only vary between individuals but are unique to each individual [83]. This is very critical as it is complicated and challenging for a model to generalize over a large group of subjects with unique patterns. Such challenges can introduce unexpected data samples into trained models [54]. These anomalies often adversely affect the feature extraction process, potentially leading to a decline in classification accuracy [24, 136]. Contemporary research [34] leverages transfer learning strategies within CNNs to generalize hand gesture data from unfamiliar subjects, drawing on the knowledge from existing datasets. Notably, inter-session datasets sometimes exhibit variability, primarily due to changes in sensor placement or the onset of fatigue. To address this, studies [187, 46] utilize domain adaptation techniques to enhance deep learning performance for motion labels across various sessions. Another significant development is proposed by Rahimian et al., using an FSL technique to address intra-session variability, i.e., unique movement repetitions within a single recording, and subject-specific variability of the sEMG data.

6.2.3 Contributions

The primary contributions of this section are:

- A novel training and testing framework is presented, leveraging D_n dataset for training, and D_f dataset for validation. The goal is to amplify the reproducibility and versatility of sEMG-based motion classification in sports science via the open-access D_n dataset.
- A new loss function, denoted as $\mathcal{L}_{\text{RCTL}}$, is introduced. This function offers advantages over the traditional triplet loss by addressing its inherent limitations. The $\mathcal{L}_{\text{RCTL}}$ loss function outperforms state-of-the-art techniques by margins of 13%, 12%, 5%, and 4% across 5, 10, 20, and 50 shot testing scenarios, respectively.
- The suggested method achieves an accuracy rate of 90% with 50-shot learning for identifying 9 distinct classes in the D_f dataset. As a reference point, the model detailed in Section 6.1.4, using a traditional training approach, achieves a 48% accuracy rate with 50 shots. This translates to a significant 41% improvement when training on a limited dataset containing 50 samples.
- Furthermore, the $\mathcal{L}_{\text{RCTL}}$ function attains accuracy levels of 65%, 78%, 83%, and 90% for 5, 10, 20, and 50 shot testing, respectively.

6.2.4 Dataset & Preprocessing

To make the twelve-channel D_n dataset consistent with the eight-channel D_f dataset, the channels in D_n are reduced to eight by selecting the most active channels for each gesture. Given a discrete signal $x_{i,j}[n]$, where i represents the channel and j indicates the gesture class, the method to determine the

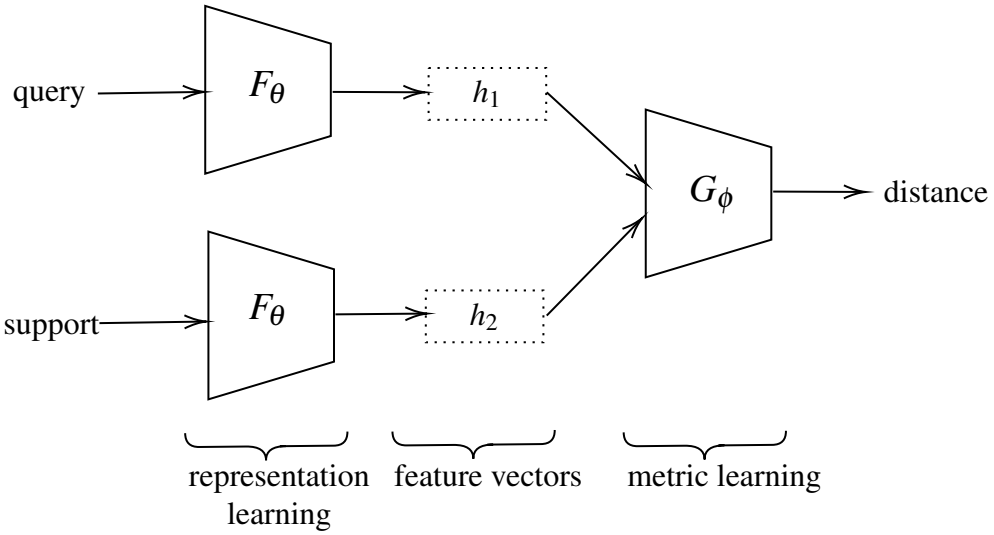


Fig. 6.5 Illustration of the few-shot metric learning (FSL) framework. The framework accepts a pair of data as input and computes their distance. Function F_θ learns representations that group same-label samples together and separates different-label samples. These representations are then passed to function G_ϕ, which calculates the distance between the paired feature vectors.

top eight channels for gesture class g_j from 49 potential gestures is

$$g_j = \text{argmax}_{j,8} \left(\sum_{n=1}^{N} \left| x_{i,j}[n] - \mu(x_{i,j}) \right| \right), \tag{6.3}$$

where N signifies the total count of time samples in each gesture recording, and $\mu(\cdot)$ denotes the mean of the signal $x_{i,j}$. For each gesture class j, the $\text{argmax}_{j,8}$ operation selects the eight channels with the highest cumulative values.

The preprocessing steps mentioned in Section 6.1.4 are followed, which incorporate Butterworth filtering, wavelet denoising, normalisation, and segmentation. Due to variations in sampling rates, data in D_f undergoes segmentation into windows of size 200, while D_n is segmented into windows of size 400. This ensures that both datasets have a consistent window duration of 200 ms.

6.2.5 Few-Shot Metric Learning

A few-shot metric learning (FSL) approach is employed to classify tasks with limited data. In FSL, predictions for a given input are based on an episodic scheme composed of a query and a support set. The query, denoted by $\tilde{x}$, is the input for which a prediction is sought. The support set, represented as $\mathcal{S}$, is defined by a union of sets

$$\mathcal{S} = \bigcup_{n=1}^{N} \left\{ \left(\bar{x}_{k,n}, y_n \right) \mid k = 1, \ldots, K \right\}, \tag{6.4}$$

where for each class n, K samples are included in the support set, encompassing samples from all N classes. The FSL problem becomes an N-way K-shot task when the support set $\mathcal{S}$ contains K samples

from each of the N unique classes. The aim is to identify the class of query $\tilde{x}$ using the samples in the support set. This involves calculating the pairwise distance between the query and every sample in the support set with a distance function $f(\cdot)$.

The label y_n in the support set that has the smallest average distance to the query is assigned as the predicted label $\tilde{y}$, represented as

$$\tilde{y} = y_n : \arg\min_{n \in \mathcal{N}} \frac{1}{K} \sum_{k=1}^{K} f(\tilde{x}, \bar{x}_{k,n}), \tag{6.5}$$

where $(\bar{x}_{k,n}, y_n) \in \mathcal{S}$.

In the FSL episodic testing framework, models are trained on one dataset and assessed on another, which includes a new set of classes. The Ninapro dataset D_n is selected for training due to its public availability, thereby enhancing the reproducibility of the methodology. The model's performance is then evaluated on the football dataset D_f using support sets with different shot sizes: 5, 10, 20, and 50. Crucially, the classes $\mathcal{C}$ from the support set $\mathcal{S}$ extracted from D_f are not utilised during training.

The distance function $f(\cdot)$, depicted in Figure 6.5 and defined in (6.5), comprises two primary components:

$$f(\tilde{x}, \bar{x}_{k,n}) = G_\phi \left(F_\theta(\tilde{x}), F_\theta(\bar{x}_{k,n}) \right), \tag{6.6}$$

where F_θ is a function, parameterised by θ, designed to learn a feature vector space that distinctively represents both similar and different classes of pairs. In contrast, G_ϕ, parameterised by ϕ, establishes a distance metric to compute scalar distances between paired vector representations.

The feature extractor from Section 6.1.4 serves as F_θ, omitting its final softmax layer, to yield a representation vector. Specifically, F_θ processes an input tensor of shape $(8, 200, 2)$ — corresponding to (channel, time, frequency) dimensions — and produces a feature vector of shape 64 such that

$$F_\theta : \mathbb{R}^{8 \times 200 \times 2} \mapsto \mathbb{R}^{64}. \tag{6.7}$$

The choice of G_ϕ is pivotal as it translates a pair of feature vectors, created by F_θ, into a singular distance metric. This translation is described as

$$G_\phi : \mathbb{R}^{2 \times 64} \mapsto \mathbb{R}^{1}. \tag{6.8}$$

Two novel G_ϕ functions are introduced in this study, and are benchmarked against existing best practices. The following subsections discuss the metric learning techniques and their associated loss functions in greater detail.

6.2.6 State-of-the-art Metric Learning (G_ϕ)

Modified Triplet Loss with Direction Regularisation ($\mathcal{L}_\mathbf{T}$)

In triplet loss [116], the metric function G is equal to the squared Euclidean distance, $d(\cdot)$, which measures the straight-line distance between two points in Euclidean space. For two vectors x_1 and x_2,

the distance $d(\cdot)$ is defined as

$$d(x_1, x_2) = \|x_1 - x_2\|_2^2 = \sum_{i=1}^{n} (x_1[i] - x_2[i])^2, \tag{6.9}$$

where n denotes the number of dimensions of the vectors. Within the model, the distance function in (6.6) translates to $d\left(F_\theta(\tilde{x}), F_\theta(\bar{x}_{k,n})\right)$.

Given a dataset D_n and a loss function $\mathcal{L}_\mathrm{T}(\theta; D)$ parameterized by θ, the ideal model parameters θ^* are:

$$\theta^* = \arg\min_\theta \mathcal{L}_\mathrm{T}(\theta; D). \tag{6.10}$$

Training employs batches of triplets: an anchor sample x_a, a positive sample x_p (having the same label as x_a), and a negative sample x_n (with a label different from x_a). To calculate the loss, the feature vectors for each triplet are determined as

$$\begin{aligned} x'_a &= F(x_a), \\ x'_p &= F(x_p), \\ x'_n &= F(x_n). \end{aligned} \tag{6.11}$$

The objective is to reduce the distance between the anchor and the positive sample while increasing the distance between the anchor and the negative sample. The distance between similar labels, d_pos and negative labels d_neg are calculated as the squared euclidean distance normalised by their sum, defined as

$$\begin{aligned} d_\mathrm{pos} &= \frac{\|x'_a - x'_p\|_2^2}{\|x'_a - x'_p\|_2^2 + \|x'_a - x'_n\|_2^2}, \\ d_\mathrm{neg} &= \frac{\|x'_a - x'_n\|_2^2}{\|x'_a - x'_p\|_2^2 + \|x'_a - x'_n\|_2^2}. \end{aligned} \tag{6.12}$$

The optimal configuration aims for the normalized distance of like pairs to approach zero, and distinct pairs to approximate the margin, α, typically set to 1. This leads to the loss component:

$$\|[d_\mathrm{pos}, d_\mathrm{neg} - \alpha]\|_2^2. \tag{6.13}$$

The direction regularization term, introduced in [116], ensures that the anchor-positive vector is oriented in a direction opposite to the anchor-negative vector. This is achieved using the cosine similarity between the vectors $(x'_n - x'_a)$ and $(x'_p - x'_a)$ such that

$$\mathrm{DR} = -\gamma \times \frac{(x'_n - x'_a) \cdot (x'_p - x'_a)}{\|x'_n - x'_a\|_2 \|x'_p - x'_a\|_2}, \tag{6.14}$$

where $\gamma = 0.4$ is the regularisation weight. Finally, the loss, $\mathcal{L}_\mathrm{T}$, is computed by taking the 2-norm of the concatenated normalized distances and adding the direction regularization as

$$\mathcal{L}_\mathrm{T} = \max\left(0, \|[d_\mathrm{pos}, d_\mathrm{neg} - 1]\|_2^2 + \mathrm{DR}\right). \tag{6.15}$$

Relation Network

The Relation Network, as detailed in [155], employs deep learning to determine a metric function, represented as G_ϕ. This technique incorporates two convolutional blocks followed by fully connected layers to process concatenated feature vectors, facilitating the learning of the metric function.

In the experiments of this section, various model architectures and data merging techniques are evaluated. Concatenating feature vectors and then processing them through two fully connected layers proves effective. The terminal layer integrates a sigmoid activation function. Table 6.5 itemizes specific layers and their associated output shapes. Both the third and fourth layers utilize ReLU activation and batch normalization. The concluding layer employs a sigmoid activation function. The model provides a distance score as a relational metric, adopting the mean squared error (MSE) loss as recommended in [155].

Table 6.5 Details of the deep learning model acting as G_ϕ, delineating the layer type and output shapes.

Layers	Type	Output Shape
1	Input	(2,64)
2	Concatenation	(128)
3	FC (Fully Connected)	(64)
4	FC (Fully Connected)	(32)
5	FC (Fully Connected)	(1)

6.2.7 Proposed Methods

Hybrid Loss ($\mathcal{L}_\mathbf{H}$)

Motivated by the strengths of triplet loss in guiding representation learning, denoted as F_θ, and the capability of the relation network to learn a more effective metric function, G_ϕ, than mere Euclidean distance, both approaches are integrated to achieve improved performance.

The relation network detailed in Section 6.2.6 is adopted for the metric function G_ϕ. A hybrid loss function is introduced, which consists of both the MSE loss, mainly updating the weights of the relation network G_ϕ, and the triplet loss, which predominantly influences the weights of the representation learner F_θ. This hybrid loss is expressed as

$$\mathcal{L}_\mathrm{H} = \lambda \mathcal{L}_\mathrm{T} + (1-\lambda)\mathcal{L}_\mathrm{MSE}. \tag{6.16}$$

The adaptive weight, λ, gauges the effectiveness of representation vectors in differentiating between losses. If the representation via F_θ meets the conditions of the triplet loss, then the metric function through deep learning can surpass the conventional Euclidean distance. Consequently, λ is modified based on two conditions, ensuring that the negative distance d_neg surpasses the threshold τ^- while the positive distance d_pos remains below the threshold τ^+. The conditions c_1 and c_2 are expressed as

$$\begin{aligned} c_1 &= \tau^+ - d_{\text{pos}}, \\ c_2 &= d_{\text{neg}} - \tau^-, \end{aligned} \tag{6.17}$$

where the thresholds are set at $\tau^+ = 0.02$ and $\tau^- = 0.9$, ensuring a larger normalized distance for distinct labels and a smaller one for similar labels.

The λ value is updated as

$$\lambda = \frac{1}{1 + e^{\max(\min(\eta, 10), -10)}}, \tag{6.18}$$

with

$$\eta = -10^3 \times c_1 \times c_2 \times (\text{sgn}(c_1) + \text{sgn}(c_2) - 1). \tag{6.19}$$

The function $\text{sgn}(\cdot)$ provides outputs of -1 for negative values, $+1$ for positive, and 0 for neutral ones. In the beginning, (6.18) leans towards a higher λ value. As the representation vector conforms to the conditions in (6.17), λ diminishes, giving more weight to $\mathcal{L}_{\text{MSE}}$, thus promoting the learning of an enhanced metric function relative to the Euclidean distance.

Regulated Centroid-Nested Triplet Loss ($\mathcal{L}_{\text{RCTL}}$)

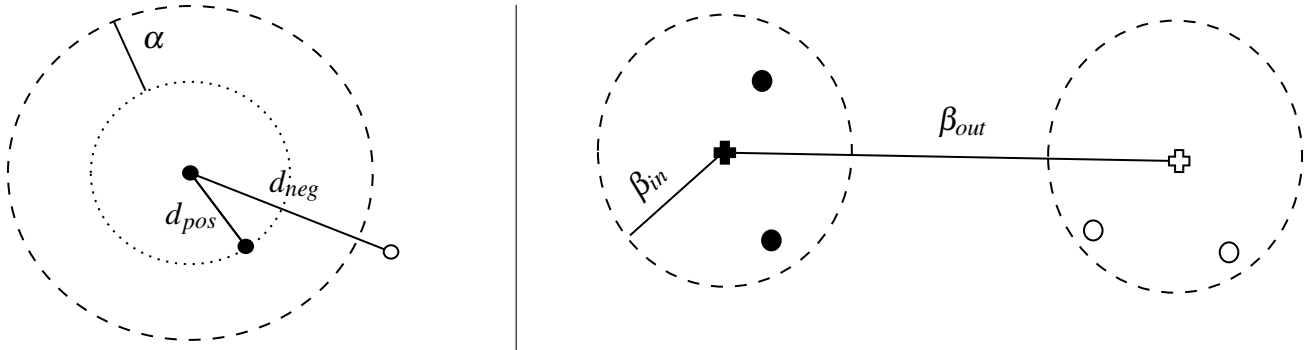
The Regulated Centroid-Nested Triplet Loss ($\mathcal{L}_{\text{RCTL}}$) refines the prevalent triplet loss to address its inherent challenges. One primary issue with the triplet loss is its dependency on extensive triplet datasets, further amplified by the pivotal role of triplet selection. Techniques such as hard negative mining have been adopted for selecting crucial triplets [66]. However, triplet loss fails to regulate the centroids' distribution. As a result, even with hard mining, intertwined ellipses can arise, leading to similarly placed centroids in the representation space. Such configurations can degrade the model's classification accuracy. The proposed $\mathcal{L}_{\text{RCTL}}$ sidesteps the need for strategies like hard-negative mining and extensive datasets, ensuring an adequate representation space. As depicted in Figure 6.6, $\mathcal{L}_{\text{RCTL}}$ enhances the triplet loss by:

- Learning label centroids as parameters instead of relying on sample means.
- Ensuring inter-centroid distances surpass the threshold β_{out}.
- Maintaining intra-label samples within the β_{in} distance of their respective centroids.

For the $\mathcal{L}_{\text{RCTL}}$, the squared Euclidean distance, denoted by $d(\cdot)$, serves as our metric function G_ϕ. The $\mathcal{L}_{\text{RCTL}}$ introduces a centroid matrix $\mathcal{C}$, consisting of randomly initialized representation vectors of dimension 64, such that $\mathcal{C} \in \mathbb{R}^{N \times 64}$. Treated as a parameter, $\mathcal{C}$ undergoes updates during backpropagation, ensuring optimal centroids for each class $n \in N$.

Consider a batch of triplets as indicated in (7.26). Let y_a and y_n represent the true labels of the anchor and negative samples, respectively. The squared Euclidean distance between their centroids is calculated as

$$d_c = ||\mathcal{C}[y_a] - \mathcal{C}[y_n]||_2^2, \tag{6.20}$$



(a) Illustration of $\mathcal{L}_{\mathrm{T}}$

(b) Illustration of $\mathcal{L}_{\mathrm{RCTL}}$

Fig. 6.6 Comparative visualizations of the conventional Triplet Loss and the proposed Regulated Centroid-Nested Triplet Loss ($\mathcal{L}_{\mathrm{RCTL}}$). (a) depicts the structure and relationships of the Modified Triplet Loss, underscoring its reliance on large triplet datasets and the importance of triplet selection. Conversely, (b) showcases the distinctive features of the $\mathcal{L}_{\mathrm{RCTL}}$, including centroid regulation and intra-class nested regions, providing a clearer and more insightful perspective on data distributions. In the visualizations, circles filled with the same color belong to the same class (e.g., all black filled circles are one class, and all white filled circles are another). Furthermore, the plus signs at the center of (b) represent the learnt centroids. Collectively, these figures highlight the advancements made to overcome the traditional triplet loss method's constraints.

where, $\mathcal{C}[y_a]$ and $\mathcal{C}[y_n]$ denote the centroids corresponding to the true labels. Next, we compute the nested distances between samples of the same label and their associated centroids, given by

$$\begin{aligned} d_a &= ||\mathcal{C}[y_a] - x'_a||_2^2, \\ d_p &= ||\mathcal{C}[y_p] - x'_p||_2^2, \\ d_n &= ||\mathcal{C}[y_n] - x'_n||_2^2, \end{aligned} \tag{6.21}$$

where, d_a, d_p, and d_n describe distances from the anchor, positive, and negative representation vectors to their respective centroids. All distances in (6.20) and (6.21) are then normalized by their sum.

To regulate the centroid distance based on the threshold β_{out}, we introduce

$$\mathcal{L}_{\mathrm{RC}} = \max(\beta_{out} - d_c, 0), \tag{6.22}$$

which penalizes the model if the centroid distances fall below the threshold $\beta_{out} = 0.8$. To ensure samples of the same class are close to their centroid, we employ

$$\begin{aligned} \mathcal{L}_{\mathrm{NC}} = &\max(d_a - \beta_{in}, 0) \\ &+ \max(d_n - \beta_{in}, 0) \\ &+ \max(d_p - \beta_{in}, 0), \end{aligned} \tag{6.23}$$

where, $\beta_{in} = 0.2$. Finally, $\mathcal{L}_{\mathrm{RCTL}}$ is defined as

$$\mathcal{L}_{\mathrm{RCTL}} = \mathcal{L}_{\mathrm{RC}} + \mathcal{L}_{\mathrm{NC}}. \tag{6.24}$$

6.2.8 Experiments

In the experiments, accuracy is utilised as the performance metric to compare all the methods outlined in this section.The training set comprises 32768 triplets sourced from D_n. Post-training, models undergo fine-tuning with support sets of sizes 5, 10, 20, and 50 shots. Only the final two convolutional layers of F_θ are updated during this phase. Adaptation to the new domain is achieved through the adaptive batch normalization method, updating only the batch normalization parameters. After fine-tuning, the model is tested on D_f using the episodic testing approach as described in Section 6.2.5. To simulate real-world conditions where support set data remains consistent due to practical constraints, the support set is kept unchanged for each episode in the test. The aforementioned training and fine-tuning process is repeated 10 times. Additionally, the support set is selected and iterated 10 times for each trained model, leading to a total of 100 iterations. The results from these iterations are subsequently averaged.

6.2.9 Results

The results in Table 6.6 display the accuracy of different models for a 9-way classification task across 5, 10, 20, and 50 shot learning scenarios. The models compared include traditional loss functions such as $\mathcal{L}_{\text{MSE}}$, $\mathcal{L}_{\text{T}}$, and a hybrid loss $\mathcal{L}_{\text{H}}$, along with our regulated centroid-nested triplet loss, $\mathcal{L}_{\text{RCTL}}$. For context, the accuracy of the traditionally trained model without FSL is also included, but only for the 50-shot scenario.

Table 6.6 Comparison of accuracy (%) among the presented models for 9-way classification using 5, 10, 20, and 50-shot learning. Traditional feed-forward training of the proposed deep learning model in Section 6.1.4 without the FSL technique is included for the 50-shot scenario only.

Model	**5-shot**	**10-shot**	**20-shot**	**50-shot**
$\mathcal{L}_{\text{MSE}}$	32	40	56	59
$\mathcal{L}_{\text{H}}$	52	61	72	73
$\mathcal{L}_{\text{T}}$	52	66	78	86
$\mathcal{L}_{\text{RCTL}}$	**65**	**78**	**83**	**90**
w/o FSL	-	-	-	49

The table clearly shows that $\mathcal{L}_{\text{RCTL}}$ consistently outperforms other models. In the 5-shot, 10-shot, 20-shot, and 50-shot learning scenarios, $\mathcal{L}_{\text{RCTL}}$ achieves accuracies of 65%, 78%, 83%, and 90% respectively. These results significantly surpass those of other models, underscoring the efficacy of $\mathcal{L}_{\text{RCTL}}$. The model trained without FSL achieves a 49% accuracy in the 50-shot scenario, further emphasizing $\mathcal{L}_{\text{RCTL}}$'s superiority.

The relation network exhibits the lowest performance, likely due to its failure to transfer knowledge in metric learning. $\mathcal{L}_{\text{H}}$, a hybrid loss combining distance-based triplet-loss with deep learning, offers an improvement over the relation network but still falls short of $\mathcal{L}_{\text{T}}$'s performance, which leverages Euclidean distance for learning and inference.

Table 6.7 provides a comprehensive analysis of average class-based precision, recall, and F1 scores for the 50-shot, 9-way classification problem across $\mathcal{L}_{\text{RCTL}}$, $\mathcal{L}_{\text{T}}$, and without FSL. Notably, $\mathcal{L}_{\text{RCTL}}$ achieves the highest F1 scores in seven of nine classes, indicating robust performance across various classification scenarios. While $\mathcal{L}_{\text{T}}$ shows competitive performance in two classes, $\mathcal{L}_{\text{RCTL}}$ consistently leads in the majority. Conversely, training without FSL generally exhibits lower performance across all classes, highlighting traditional training approaches' limitations in few-shot settings.

Table 6.7 Average class-based precision, recall, and F1 score for a 50-Shot, 9-Way classification. This table presents a comparative analysis of three different methods: $\mathcal{L}_{\text{RCTL}}$, $\mathcal{L}_{\text{T}}$, and w/o FSL. For each class (1 through 9), the table displays the precision, recall, and F1 score metrics, with the highest F1 score for each class highlighted in bold.

	$\mathcal{L}_{\text{RCTL}}$			$\mathcal{L}_{\text{T}}$			w/o FSL		
	Precision	**Recall**	**F1 score**	**Precision**	**Recall**	**F1-Score**	**Precision**	**Recall**	**F1-Score**
Class 1	0.90	0.79	**0.84**	0.79	0.77	0.78	0.44	0.26	0.33
Class 2	0.98	0.96	**0.97**	0.98	0.92	0.95	0.64	0.70	0.67
Class 3	0.93	0.92	**0.93**	0.85	0.91	0.88	0.42	0.33	0.37
Class 4	0.89	1.00	0.94	0.98	0.97	**0.98**	0.71	0.88	0.78
Class 5	0.91	1.00	0.96	0.99	1.00	**1.00**	0.67	0.64	0.66
Class 6	0.85	0.90	**0.87**	0.68	0.80	0.73	0.38	0.47	0.42
Class 7	0.84	0.76	**0.80**	0.82	0.70	0.76	0.31	0.45	0.37
Class 8	0.86	0.84	**0.85**	0.85	0.84	0.84	0.33	0.23	0.27
Class 9	0.91	0.92	**0.92**	0.80	0.87	0.83	0.57	0.43	0.49

6.2.10 Discussions

The enhanced efficacy of the proposed $\mathcal{L}_{\text{RCTL}}$ model over traditional triplet loss methods, as demonstrated in our research, underscores the effectiveness of the regulated centroid-nested triplet loss approach. This methodology adeptly manages data distributions, promoting dense clustering within classes while ensuring clear separation between them. The success of $\mathcal{L}_{\text{RCTL}}$ in sEMG-based motion classification, particularly in few-shot scenarios, is attributable to several pivotal features.

Initially, $\mathcal{L}_{\text{RCTL}}$ introduces a novel mechanism for learning class centroids dynamically in each iteration, significantly enhancing the model's representational capacity. Contrasting with traditional methods where centroids are calculated as mean feature vectors, in $\mathcal{L}_{\text{RCTL}}$, centroids are treated as parameters updated iteratively. This dynamic learning approach yields a more accurate and representative feature space.

Furthermore, $\mathcal{L}_{\text{RCTL}}$ regulates the distance between class centroids through a pre-specified normalized distance, a critical feature for maintaining clear class separation. This regulation penalizes the model when different class clusters converge, a significant improvement over traditional triplet loss

methods reliant on triplet sampling. The regulated centroids in $\mathcal{L}_{\text{RCTL}}$ ensure consistent separation between classes, independent of triplet configuration in the data.

Moreover, the loss function in $\mathcal{L}_{\text{RCTL}}$ enforces proximity within the same class, mandating that vectors adhere to a predetermined normalized Euclidean distance, thereby fostering intra-class compactness. This approach is particularly advantageous in sEMG metric learning, addressing issues where traditional triplet or contrastive-based loss methods may lead to overly dispersed or intertwined feature vectors, which can reduce prediction performance. $\mathcal{L}_{\text{RCTL}}$ effectively penalizes scenarios where centroids of different classes are in close proximity.

6.3 Conclusion

In this chapter, a deep learning model for motion classification is introduced, which shows significant advancements over contemporary models on several datasets, including D_n and $D_{n,b}$. The model registers remarkable accuracies of 95.30% and 98.80% for D_n and $D_{n,b}$, respectively. This outperformance is evident when comparing to the leading contemporary models, with improvements of 3.29% and 8.66%. The proposed model's primary strength is its adaptability across a variety of datasets, achieved without the necessity for parameter tuning. This capability is evidenced by the model's consistent performance, notably achieving a 96.30% accuracy on the D_f dataset, with no alterations in its parameters or layer structure. The only modification required pertains to the size of the final output layer, which is adjusted to correspond with the number of classes in each specific dataset. The model's versatility across different datasets is largely due to its strategic utilization of various kernel sizes and depths, coupled with the integration of an attention module, which together contribute to its flexible and robust performance.

The main limitation of the proposed deep learning model is its dependence on extensive training data, a characteristic inherent to traditional deep learning models which require large datasets for optimal performance. In sports science applications, this issue is compounded by the often limited availability of datasets. To address this challenge, the study implements FSL for motion classification within sports science, effectively overcoming the constraints of dataset size. A significant advancement in this study is the introduction of the $\mathcal{L}_{\text{RCTL}}$ model, which outperforms traditional triplet loss methods. The $\mathcal{L}_{\text{RCTL}}$ achieves accuracies of 65%, 78%, 83%, and 90% in 5, 10, 20, and 50 shot testing scenarios, respectively. These results represent improvements of 13%, 12%, 5%, and 4% over existing methodologies. Moreover, the $\mathcal{L}_{\text{RCTL}}$ model surpasses the performance of conventional deep learning training, not utilizing FSL, by a notable 41%. This underscores the effectiveness and potential applicability of the $\mathcal{L}_{\text{RCTL}}$ model in sports science scenarios, where limited dataset sizes are a common challenge.

Chapter 7

Hamstring, Adductor, Soleus Injury (HASI) Prediction

7.1 Introduction

This chapter clarifies the final block of the presented framework. Preceding blocks supply this framework with the MIA denoised segment of drills and relevant drill class information. The primary objective of this chapter is to forecast whether an athlete faces a high or low risk of injury. This issue is structured as a binary classification problem where 1 denotes a high risk of injury, and 0 indicates a low risk.

The central hypothesis of this thesis posits that athletes, when fatigued or undergoing excessive strain due to significant external load, tend to utilise their muscles differently. This deviates from their established baseline of prior normal, i.e., healthy, measurements, which did not lead to injuries. Hence, injury prevention aligns with anomaly detection, where injuries are viewed as data variations or anomalies—diverging from typical sEMG signals. As detailed in Chapter 1 and Chapter 2, indirect injuries predominantly arise from overload and fatigue, peaking when athletes participate in an official match day (MD). Consequently, the sEMG screenings in this research focus on MD+1 measurements to gauge the deviation from athletes' MD+1 baseline—cluster of earlier screenings. In this context, D_f is employed as the primary dataset, and all suggested models undergo testing on this dataset.

Two distinct approaches for HASI risk prediction are proposed in this chapter. The first approach leverages manual feature selection to identify salient features indicative of risk factors from the sEMG signal, employing signal processing algorithms. Once these features are extracted, the Gaussian distribution model is applied to determine if a particular sEMG data sample indicates injury.

The subsequent approach engages feature learning through deep metric learning. Rather than relying on manually selected signal processing-based features, this technique employs a deep neural network to discern features. It constructs a representation space where safe (low risk measurements) cluster closely to their centroid. In contrast, risky (high risk measurements) exhibit greater distances from the normal measurement cluster.

Throughout this chapter, the term injury refers to the specific HASI as outlined in Chapter 2. The proposed methodologies undergo testing and validation using the football dataset, D_f. This dataset comprises MD+1 screenings of athletes when fit and prior to incurring injuries, making it an apt fit to evaluate the thesis's injury prediction hypothesis before actual injury onset.

7.2 Review of sEMG Responses to Muscular Injuries

sEMG amplitude, inter-muscle ratio of sEMG amplitude [21, 126, 133], the ratio of one hamstring muscle compared to the total sEMG amplitude of the hamstrings [133, 18], integrated sEMG [127], median power frequency of sEMG [126], between-limb differences in sEMG amplitude [18], and sEMG onset timing of various muscles [153] are recognized as the primary muscle injury indicator features in the literature.

Sole et al. demonstrate that the sEMG onsets of biceps femoris and medial hamstrings are significantly earlier for subjects with hamstring injuries compared to a control group without injuries, indicating altered neuromuscular control [153]. A recent systematic review, encompassing 12 papers and involving a total of 220 subjects, finds moderate to limited evidence that individuals with lower limb muscle strain injuries show deficits in muscle activity and activation, particularly in previously injured biceps femoris during maximal eccentric contractions. Notable differences are observed in inter-muscle sEMG ratios of biceps femoris to medial hamstring and biceps femoris to gluteus maximus, as well as between-limb differences in sEMG amplitude in limbs with injuries. This evidence suggests that hamstring strain injuries lead to persistent muscle activity disparities when comparing injured limbs to uninjured sides and controls during eccentric contractions [133].

The literature often reveals significant limitations concerning the datasets used. Commonly, studies compare the injured leg's sEMG data contralaterally, that is, with the same muscle group of the uninjured leg [133]. This approach is limiting as it may not account for pre-existing imbalances in muscle sEMG data, potentially due to factors like footedness or specific physiological demands of an athlete's position. Consequently, observed discrepancies in sEMG features might reflect pre-existing physiological imbalances rather than the effects of injury. For instance, the left biceps femoris of an athlete might inherently exhibit lower sEMG amplitude compared to the right biceps femoris, suggesting that such differences do not necessarily stem from the injury itself.

Alternatively, some studies contrast injured subjects with a control group of uninjured individuals [133, 153]. This approach introduces significant variability, as each individual exhibits unique recruitment patterns and sEMG responses [83]. Consequently, this methodological strategy raises concerns regarding the accuracy of attributing observed differences in sEMG features solely to injuries. It is important to consider that sEMG responses are inherently subject-dependent and can vary considerably from one individual to another [83]. Therefore, any observed discrepancies in sEMG data between the two groups might reflect inherent physiological differences among individuals rather than the effects of injuries.

In contrast, this thesis involves the systematic collection of comprehensive, long-term data from a diverse cohort of athletes, incorporating multiple sEMG screenings of the same drills conducted both

before and after injury incidents. This methodology enables a more precise evaluation of how injuries affect individual muscles and the overall sEMG response. It also critically examines sEMG data for the same athlete across different states—before and after injury—thus providing a direct comparison that is not limited by the typical methods used in other studies, which often involve comparisons either with the athlete's uninjured contralateral muscle group or with a control group of uninjured athletes.

7.3 Selection of sEMG-based Features As Risk Factors

The features discussed in this section are derived using signal processing methods applied to the sEMG streams. These features play a crucial role in the anomaly detection technique described in Section 7.4.

7.3.1 Cumulative Recruitment (CR)

Cumulative Recruitment (CR)% offers insight into muscle activation by evaluating the relative contribution of each muscle in comparison to the total activation of all monitored muscles. CR% quantifies the percentage contribution of a specific muscle's activation against the combined activation of all observed muscles during a given activity.

This metric clarifies the role of individual muscles within the broader muscular network under observation. For example, a notably high CR% for a specific muscle during a movement implies its significant role in that action relative to other muscles.

In the sports domain, where the relationship and contributions between muscles are crucial, CR% differs from conventional metrics like Maximum Voluntary Isometric Contraction (MVIC). Whereas MVIC measures the peak force a muscle can produce individually, CR% evaluates the muscle's contribution in a collective muscular action. This difference becomes especially evident in sports scenarios, where muscle coordination often takes precedence over the strength of an individual muscle.

For a given player's activation associated with a recognised drill, CR is determined using a sliding window technique with a window length of 128 and an overlap of 96. Subsequently, the sample entropy (SampEn) of each window is computed. The specifics of the SampEn calculation are discussed in Chapter 5. After the SampEn transformation, an 8-channel vector for SampEn representation is formed, with each channel representing a muscle group. This vector is denoted as $S[m,n]$. The collective activation contribution of SampEn values is then computed as

$$\Sigma_m = \sum_{n=0}^{W-1} S[m,n], \tag{7.1}$$

$$\Sigma_{\text{all}} = \sum_{m=1}^{8} \Sigma_m[m]. \tag{7.2}$$

In the equations above, m denotes the channel number, which ranges from 1 to 8. The variable n represents each window, spanning from 0 to $W-1$, where W specifies the total number of windows. The term Σ_m characterises the cumulative activation for the muscle indexed by m over all the windows.

Conversely, Σ_{all} captures the entire activation across all muscle groups and windows.

$$CR[m] = \frac{\sum_{n=0}^{W-1} S[m,n]}{\Sigma_{\text{all}}} \times 100 \tag{7.3}$$

As shown in (7.3), $CR[m]$ indicates the Cumulative Recruitment percentage for muscle m. The numerator sums the activation of muscle m over all windows. Conversely, the denominator, Σ_{all}, represents the combined activation across all muscles and windows. By computing the ratio of the muscle-specific sum to the total activation and multiplying by 100, this metric elucidates the relative contribution of each muscle, shown as a percentage.

7.3.2 Compensation

Compensation refers to the phenomenon where a reduction in the activity of one muscle group in a leg is offset by increased use of other muscle groups in the same leg. This concept plays a crucial role in injury prediction. Typically, athletes tend to amplify compensation notably under load and before an injury. Such patterns can be regularly monitored, especially for MD+1 evaluations.

Compensation analysis begins by first determining the bilateral imbalance of cumulative sums, Σ_g, for each muscle group. This step involves quantifying the difference between the right and left activations for corresponding muscles. The bilateral imbalance, represented as Δ_g, for each muscle group can be expressed as

$$\Delta_g = \Sigma_{g+4} - \Sigma_g, \quad \text{for } g = 1,2,3,4. \tag{7.4}$$

In this configuration, positive values of Δ_g denote a higher activation on the right side, whilst negative values indicate more pronounced activation on the left side.

Next, the imbalance array is sorted in ascending order as

$$\Delta_{\text{sorted}}[g] = \text{sort}(\Delta_g), \quad \text{for } g = 1,2,3,4, \tag{7.5}$$

where, the operation $\text{sort}(\cdot)$ organises the Δ values in ascending order. Consequently, $\Delta_{\text{sorted}}[1]$ represents the muscle group with the smallest bilateral imbalance, while $\Delta_{\text{sorted}}[4]$ will indicate the group with the most pronounced imbalance.

Compensation is computed as the aggregate difference between consecutive elements in the sorted imbalance array Δ_{sorted}. Formally, it is defined by:

$$\text{compensation} = \sum_{g=0}^{N-2} \left(\Delta_{\text{sorted}}[g+1] - \Delta_{\text{sorted}}[g]\right), \tag{7.6}$$

where N denotes the length of the array Δ_{sorted}. This mathematical representation captures the accumulative effect of differences across the sorted muscle imbalances, thus quantifying the overall compensation.

Consider a comparison between athletes categorised as risky and safe during hamstring claw exercises. Figure 7.1 presents the CR% plots for two athletes performing this exercise. In plot (a),

the athlete consistently favours the right leg across all muscle groups. This pattern results in a low compensation value of 2.49%, categorising the athlete as safe. Conversely, in plot (b), the athlete compensates for the sEMG activity reduction in the right hamstrings by highly activating the right adductor longus (AL). This leads to a significant compensation score of 33.78%, classifying the athlete as risky due to the elevated compensation value.

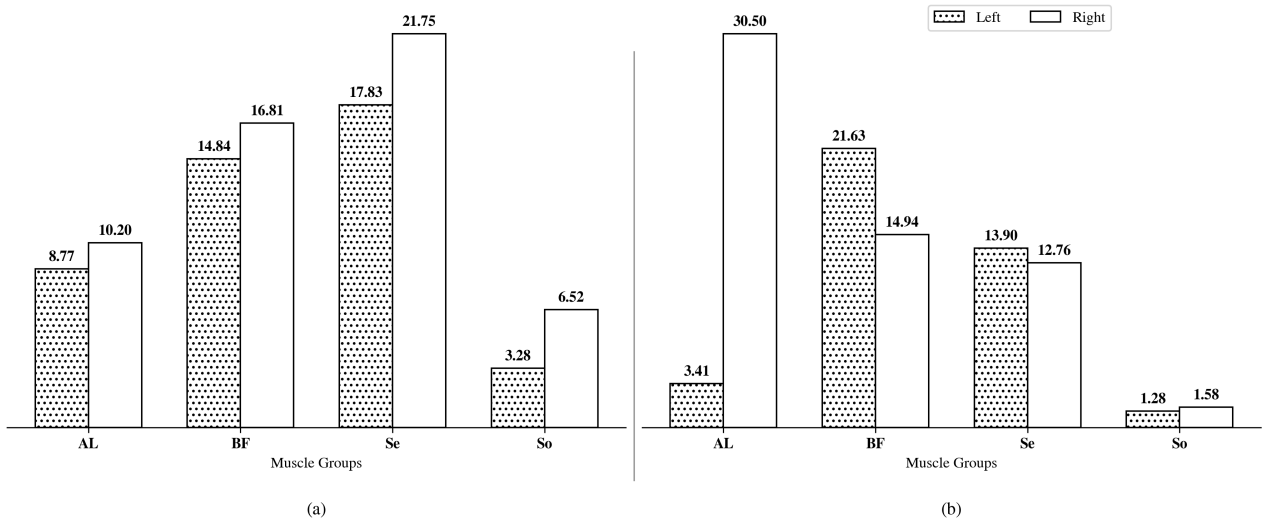


Fig. 7.1 This illustration presents the cumulative recruitment percentages for muscle groups: biceps femoris (BF), semitendinosus (Se), adductor longus (AL), and soleus (So) across two athletes performing the hamstring claw exercise. In plot (a), the athlete exhibits low compensation, consistently showing a right-dominant asymmetry. Conversely, plot (b) showcases an athlete with high compensation, where the reduced activation in the right hamstring is offset by heightened activation in the left hamstring.

7.3.3 Bilateral Similarity (BS)

The bilateral similarity evaluates the *CR%* for each leg independently, rather than assessing all muscles collectively. Consequently, the *CR%* values for each leg sum up to 100%. This is mathematically expressed as:

$$CR_{\text{left}}[m] = \frac{\sum_{n=0}^{W-1} S_{\text{left}}[m,n]}{\Sigma_{\text{all, left}}} \times 100, \tag{7.7}$$

$$CR_{\text{right}}[m] = \frac{\sum_{n=0}^{W-1} S_{\text{right}}[m,n]}{\Sigma_{\text{all, right}}} \times 100, \tag{7.8}$$

where $\Sigma_{\text{all, left}}$ and $\Sigma_{\text{all, right}}$ denote the total activations for all left and right muscles, respectively. Specifically, in this chapter, there are four muscles attributed to the left leg and another set of four muscles for the right leg. Hence, the cumulative activations for the left and right are computed considering their respective muscle groups exclusively. Then, the bilateral similarity, *BS*, is calculated using the cosine similarity between the cumulative distributions of the right and left legs, which is expressed as

$$BS = \frac{CR_{\text{left}} \cdot CR_{\text{right}}}{||CR_{\text{left}}||_2 \times ||CR_{\text{right}}||_2} \tag{7.9}$$

where CR_{left} and CR_{right} are vectors representing the cumulative distributions for the left and right legs, respectively. The dot product between the vectors is divided by the product of their magnitudes to obtain the cosine similarity.

As can be seen in Figure 7.2, bilateral similarity measures how closely the muscle recruitment distribution in the right leg aligns with that of the left leg, given in percentage terms. A 100% value means both legs share the same recruitment patterns, whereas a 0% value signifies entirely different recruitment patterns.

It is essential to distinguish between bilateral similarity and bilateral imbalance. A bilateral similarity of 100% does not imply a complete balance since this metric focuses on the similarity of distribution patterns rather than the intensity or power of recruitment. Therefore, even if muscles show the same recruitment patterns, they might still have pronounced imbalances, and the opposite may also be true.

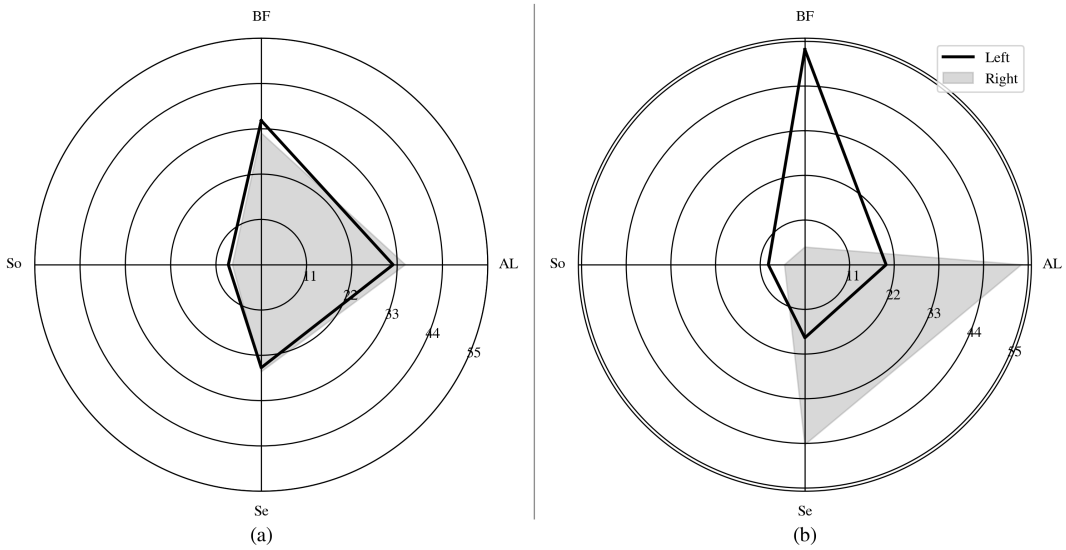


Fig. 7.2 This figure delineates the cumulative recruitment percentages of left and right muscle groups—biceps femoris (BF), semitendinosus (Se), adductor longus (AL), and soleus (So)—for two athletes undertaking the hip flexion exercise. In subplot (a), the athlete displays a pronounced bilateral similarity of 99%, indicative of a substantial overlap in the regions. Conversely, subplot (b) illustrates an athlete with a bilateral similarity of 51%, highlighting the disparities in cumulative recruitment between the left and right leg, suggesting the utilization of distinct muscles in each limb.

7.3.4 Target Muscle Imbalance (Im)

This feature quantifies the imbalance between left and right target muscles, which are specific to different drill types. Each drill type requires a more substantial activation of certain muscles compared to others. Details about the specific target muscles for each drill type can be found in Chapter 3. The imbalance is calculated using the following formula

$$\mathrm{Im}[m] = \frac{\sum_{n=0}^{W-1}(S_{\mathrm{left}}[m,n]) - \sum_{n=0}^{W-1}(S_{\mathrm{right}}[m,n])}{\sum_{n=0}^{W-1}(S_{\mathrm{left}}[m,n]) + \sum_{n=0}^{W-1}(S_{\mathrm{right}}[m,n])} \times 100. \tag{7.10}$$

7.3.5 Waveform Peak Detection

The athletes realise each drill with several repititions, which is detailed in Chapter 3. Each repetition of muscle activation starts from low sEMG activation or no activation, goes to the peak and then returns to the low activation again, this pattern resembles a bell shaped curve. Thus, in each repetition, via the peak detection algorithm, the peaks of each repetition can be identified.

Given a data sequence $S[n]$, where n represents the n^{th} sample, the process of peak detection is vital for discerning local maxima that surpass predefined prominence and temporal distance. This subsection elaborates on the methodology, using mathematical representations to convey the underlying principles.

A key condition for any point $S[n]$ to be considered as a candidate peak is its qualification as a local maximum within a designated window of size w. Mathematically, this can be articulated as

$$S[n] \geq S[k] \quad \forall \quad k \in \{n-w, \ldots, n+w\} \setminus \{n\}. \tag{7.11}$$

This constraint ensures that within the specified window size, $S[n]$ is the highest value.

Prominence plays a pivotal role in evaluating the relative height of a peak concerning its neighboring points. To elaborate, the prominence of a peak gauges the difference between the peak's height and the height of the highest contour line encompassing it but excluding any higher peaks. Let $\mathrm{base}[n]$ represent the highest value on either side of the peak where $\mathrm{base}[n]$ is lesser than $S[n]$ and all intermediate points between n and $\mathrm{base}[n]$ are also less than or equal to $\mathrm{base}[n]$. The prominence, denoted as $\mathrm{prom}[n]$, can be expressed as

$$\mathrm{prom}[n] = S[n] - \mathrm{base}[n] \tag{7.12}$$

In this study, the prominence threshold, τ_{prom}, is determined by the difference between the 90th percentile and the 40th percentile of the sequence, represented as

$$\tau_{\mathrm{prom}} = P_{90}(S) - P_{10}(S), \tag{7.13}$$

where $P_{90}(S)$ and $P_{40}(S)$ are the 90^{th} and 10^{th} percentiles of the sequence S, respectively. Only peaks with a prominence that exceeds the specified prominence threshold as

$$\mathrm{prom}[n] > \tau_{\mathrm{prom}}, \tag{7.14}$$

are considered valid.

To avoid any misinterpretations resulting from peaks positioned in close proximity, a temporal distance constraint is set. For two peaks at $S[i]$ and $S[j]$ (with $i < j$), they must adhere to

$$j - i > \tau_{\mathrm{distance}}, \tag{7.15}$$

where τ_{distance} is the distance threshold.

Algorithm 2 Peak Detection for sEMG Data

Require: $S[m,n]$ - sEMG data for each muscle index m, with W values.

procedure INITIALPEAKIDENTIFICATION(S,m)
    $upper\_bound \leftarrow$ 90th percentile of $S[m,n]$
    $lower\_bound \leftarrow$ 40th percentile of $S[m,n]$
    $prominence \leftarrow upper\_bound - lower\_bound$
    $peaks \leftarrow$ array of peaks identified in $S[m,n]$ based on *prominence*.
end procedure

procedure DETERMINETEMPORALSPACING(S,m)
    $peak\_indices \leftarrow$ array of indices where peaks occur in $S[m,n]$.
    $spacing\_arr \leftarrow$ compute difference between consecutive elements in $peak\_indices$.
end procedure

procedure COMPUTEMINIMUMSPACING($spacing\_arr$)
    $Q1\_spacing \leftarrow$ 25th percentile value from $spacing\_arr$. **return** $Q1\_spacing$
end procedure

procedure REFINEDPEAKDETECTION($S,m,Q1\_spacing$)
    Identify peaks in $S[m,n]$ based on both *prominence* and spacing $\geq Q1\_spacing$.
end procedure

procedure PEAKDETECTIONMETHODOLOGY(S,m)
    INITIALPEAKIDENTIFICATION(S, m)
    DETERMINETEMPORALSPACING(S, m)
    $Q1\_spacing \leftarrow$ COMPUTEMINIMUMSPACING($spacing\_arr$)
    REFINEDPEAKDETECTION(S, m, Q1_spacing)
    return Peaks in $S[m,n]$ representing genuine muscle contractions.
end procedure

In Algorithm 2, the algorithmic representation of the proposed peak detection is given. Building on the foundational principles of peak detection, this method presents a refined approach tailored to the nuances of sEMG data from various muscle groups. In the initial phase, each muscle group is analysed independently. For each muscle group m, peaks are pinpointed using the prominence criterion without factoring in the temporal distance between them. After determining the peaks for every muscle group, the temporal spacing between these peaks is assessed, highlighting the periodic nature of muscle activations during repetitions. To achieve this, the discrete derivative of the array of peak indices for each muscle group m is computed. This step measures the distance between successive peaks, shedding light on the activation's rhythmic pattern. To determine a representative minimum spacing between

peaks, the 25^{th} percentile of the peak distances array, obtained from the derivative, is evaluated. This computation offers a resilient estimate that overlooks sporadic aberrations from irregular activations, ensuring the derived value genuinely represents the prevalent rhythmic pattern. With this data at hand, the peak detection process is performed once more for each muscle group, this time incorporating the earlier computed distance threshold. This refinement ensures that the recognised peaks are not just of significant amplitude but also suitably spaced, capturing the authentic nature of muscle activations during repetitions.

This technique not only assures precise peak detection but is also flexible to the distinct temporal patterns present in the sEMG data from varied muscles. By evaluating both the prominence and temporal spacing between peaks, the algorithm accurately detects activations representing genuine muscle contractions during drills.

7.3.6 Explosiveness (V/s)

The measure of muscular explosiveness provides insights into the rapidity and intensity of muscle contractions. This is quantified through the impulsiveness of the muscle's contraction, particularly by assessing the slope of the signal from its initiation to its peak. Let $S[m,n]$ represent the peak of the m^{th} muscle at the n^{th} index, and $S[m,n-k]$ be the baseline value where the activation commences. The explosiveness, denoted as exp, can be expressed as:

$$\text{exp} = \frac{S[m,n] - S[m,n-k]}{k}. \tag{7.16}$$

For comprehensive analysis, this calculation is applied across all detected peaks, subsequently determining the median explosiveness value. Notably, in this research, emphasis is placed on the median explosiveness of only the target muscle groups, underscoring the significance of specific drill information.

7.3.7 Repeatability of sEMG Measurements Across Various Screenings

The variation of sEMG data across different sessions may be high due to physiological changes (such as variations in fatigue) of the subject being recorded and the electrode placement. The placement of the electrode across different areas of the muscle can yield significantly different results. Studies employing high-density surface EMG have shown clear differences in activation in the hamstrings across different exercises [76].

The data collected in this thesis includes multiple measurements of the same activities across different days. Therefore, it is important to conduct a repeatability check to understand the variation caused by electrode placement and physiological state changes. As a repeatability check, the coefficient of variation (CV) is selected, which is defined mathematically as

$$\text{CV} = \left(\frac{\sigma}{\mu}\right) \times 100, \tag{7.17}$$

where σ is the standard deviation and μ is the mean of the distribution.

The peak median of each muscle is selected as the feature since it reflects each muscle's activation magnitude. Each athlete may have unique activation patterns, and an athlete's activation pattern is also dependent on the drill being performed. Therefore, the coefficient of variation is applied to each feature of each athlete for each drill separately, and then the median per feature is calculated, as displayed in Table 7.1. Additionally, since the dataset D_f includes recordings with pre-injury states and some highly fatigued versions of athletes, which create high variability and distance from the mean, those with risky measurements and highly fatigued measurements are removed to only consider the day-to-day variations when the athlete is not in a risky physiological state.

Table 7.1 Median Coefficient of Variation for Peak Median Values Across Various Muscle Groups

Feature	**Coefficient of Variation (%)**
Peak Median BF-L	26.09
Peak Median BF-R	24.31
Peak Median S-L	26.55
Peak Median S-R	30.44
Peak Median AL-L	50.43
Peak Median AL-R	45.00
Peak Median So-L	21.42
Peak Median So-R	26.94

As seen in Table 7.1, the coefficient of variation is marginally above 20% and below 30% for the muscles biceps femoris (BF), semitendinosus (S), and soleus (So), which indicates that the repeatability in our measurements is valid considering sEMG standards and varying fatigue levels. The adductor longus (AL), on the other hand, has high variability with the coefficient of variation value being above 45%. This is understandable as the adductor muscle is observed as a major muscle for compensation, often recruited to compensate for the sEMG activity reduction in the hamstrings or imbalances caused by fatigue from excessive changes in direction.

7.3.8 Validation of Compensation & Bilateral Similarity Metrics

An experiment was conducted to evaluate the effectiveness of compensation and bilateral similarity metrics in collaboration with Manchester United football club for the 2022/2023 pre-season. In this experiment, 24 healthy male elite football players were engaged during pre-season screening. Muscle activity, as captured using EMG, was bilaterally measured for the adductor longus, biceps femoris long head, and semitendinosus during the hamstring claw (prone) exercise on a force frame. Each athlete performed three repetitions of maximal isometric contractions for each leg individually. The average of the peak portions of the contractions was used in subsequent calculations. To integrate data from isometric testing with EMG, the imbalance (in %) between left and right torque values from isometric

testing and recruitment anomaly (in %) from EMG measurements was considered. The recruitment anomaly is determined as the cosine distance between a vector of compensation and bilateral similarity values for each player compared to a reference vector. In this reference vector, bilateral similarity is defined at 100% as ideal, and compensation is set at 10%, approximating the average for non-injured EPL athletes measured to date.

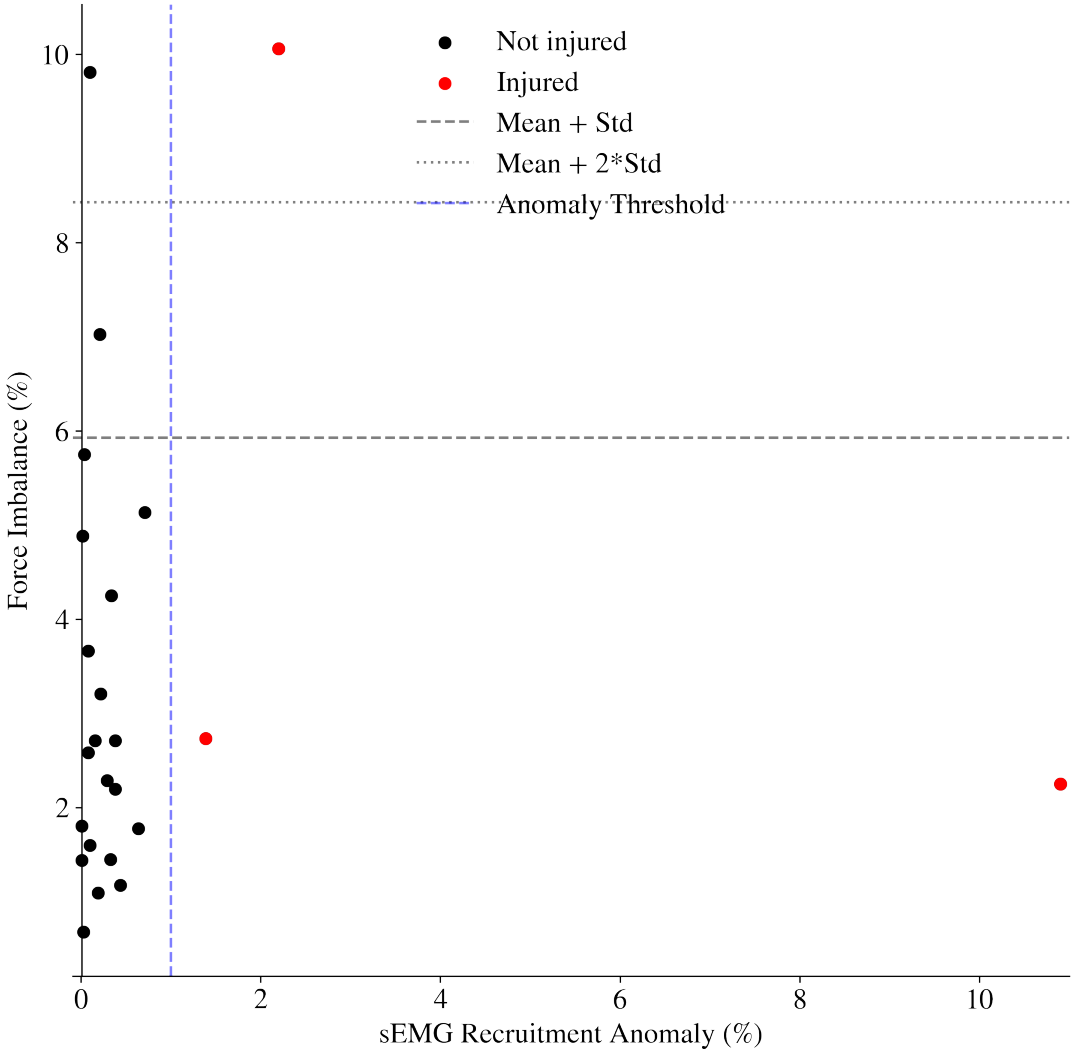


Fig. 7.3 This figure illustrates the juxtaposition of external load force imbalance percentage (y-axis) against the internal load sEMG recruitment anomaly percentage (x-axis) during the pre-season screening for Manchester United Football Club. Red markers denote players who sustained hamstring muscle injuries within the first three months of the 2022/2023 season, whereas black markers represent players without any injuries during the same timeframe. The graph suggests the potential utility of internal sEMG data and recruitment anomaly metrics as risk indicators in pre-season evaluations, complementing external load metrics.

Of the 24 assessed players, three sustained hamstring-related injuries in the three months post pre-season screening. All injured players exhibited an EMG recruitment anomaly exceeding 1%, with specific values at 1.39%, 2.23%, and 10.92%. The median recruitment anomaly percentage for uninjured players stands at 0.19%, with a peak value of 0.71%. Consequently, a 1% recruitment anomaly was chosen as the defining threshold to distinguish between injured and uninjured player distributions. The isometric force imbalances for the injured players were 2.76%, 10.10%, and 2.23%. For uninjured players, the median force imbalance was 2.64%, with the highest recorded value being 10.06%.

Both the force imbalance and the sEMG recruitment anomaly metrics are depicted in Figure 7.3. In this representation, injured players are highlighted in red and non-injured in black, with a clear separation at the 1% threshold. The graphic also underscores that the force imbalance data alone does not allow for a linear separation between injured and non-injured players. This study underscores the significance of the introduced metrics of compensation and bilateral similarity presented in this chapter.

The combined consideration of external and internal loads introduces a nuanced perspective to hamstring injury prevention. Isometric strength tests are especially pertinent as they enable maximal isometric contractions with precise control over hip/knee angles and yield external load data such as force readings. Nonetheless, this chapter emphasises that without insight into the internal load during such isometric tests, injury prediction remains inadequate. sEMG sheds light on the internal load, revealing how muscles are internally activated. Moreover, the recruitment anomaly emerges as a promising marker for injury prevention during pre-season screenings, encapsulating insights on bilateral recruitment patterns and compensation.

7.4 HASI Prediction via Anomaly Detection with Independent Univariate Gaussian Distributions (UGAD)

The method, termed Univariate Gaussian Anomaly Detection (UGAD) for this thesis, is a specialized anomaly detection approach rooted in statistical analysis. It is premised on the assumption that the features of sEMG data, representing 'normal' instances, adhere to a Gaussian distribution. Anomalies are identified as outliers that significantly deviate from this distribution, often characterized by their substantial distance from the mean. The fundamental objective of this approach is to quantify the deviation of each data point from the Gaussian norm, facilitating its classification as either 'normal' or 'anomaly'.

In this approach, features detailed in Section 7.3 are employed. Importantly, this methodology operates under the assumption that each feature is independent. This assumption simplifies the model and focuses on the main effects of individual features, despite not fully aligning with the inherent dependencies among muscles observed in sEMG data. Typically, muscle activations are interdependent because muscles often work together to compensate for each other's functions and to collectively support the body's ability to generate external force and perform activities, as discussed in Section 7.3.2.

However, the decision to treat each feature independently is deliberate, allowing the model to focus more on analyzing independent channels and reducing model complexity. Within the scope of this chapter, the model is trained primarily using 'normal' data, specifically from screenings of fit, low-risk players.

Consider the training dataset $\mathbf{X} = \{x^{(1)}, x^{(2)}, \ldots, x^{(m)}\}$, where each data instance $x^{(i)}$ is represented as $x^{(i)}[j] = \{x^{(i)}[1], x^{(i)}[2], \ldots, x^{(i)}[\mathcal{N}]\}$ with j ranging from 1 to $\mathcal{N}$. Key features include compensation, bilateral similarity, explosiveness of target muscles (left and right), target muscle imbalance, and the median of peak values for each muscle. Further details on these features can be found in Section 7.3. Given that there are 13 features for each sEMG screening data sample, it is established that $\mathcal{N} = 13$. The model hypothesises that each feature j of the data instances conforms to a Gaussian distribution. The probability density function for this Gaussian distribution is given by

$$p(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{7.18}$$

where, for given values of μ and σ^2, this equation provides the likelihood or density of the value x occurring relative to other possible values in the distribution.

For every feature $j \in [1, \mathcal{N}]$, parameters μ_j (mean) and σ_j^2 (variance) can be estimated using

$$\mu_j = \frac{1}{m} \sum_{i=1}^{m} x^{(i)[j]}, \tag{7.19}$$

$$\sigma_j^2 = \frac{1}{m} \sum_{i=1}^{m} (x^{(i)[j]} - \mu_j)^2, \tag{7.20}$$

where m is the total number of data instances in the training set.

To determine if a query data instance $x^{(q)}$ is an anomaly, instead of directly computing the product of the probabilities for each feature's Gaussian distribution, the sum of the log probabilities is often used. This approach, using the logarithm, ensures numerical stability and is computationally efficient. The equation becomes

$$\log(p(x^{(q)})) = \sum_{j=1}^{\mathcal{N}} \log(p(x^{(q)}[j]; \mu_j, \sigma_j^2)). \tag{7.21}$$

Given a validation set $\mathbf{V} = \{v^{(1)}, v^{(2)}, \ldots, v^{(l)}\}$, the threshold ε for anomaly detection can be determined, by computing (7.22) for each data instance $v^{(i)}$ in $\mathbf{V}$.

$$\log(p(v^{(i)})) = \sum_{j=1}^{\mathcal{N}} \log(p(v^{(i)}[j]; \mu_j, \sigma_j^2)) \tag{7.22}$$

Performance evaluation of various ε values on $\mathbf{V}$ leads to the selection of an ε that optimizes the F1-score. This optimal threshold is represented as ε^* and is defined by

$$\varepsilon^* = \arg\max_{\varepsilon} F1(\varepsilon, \mathbf{V}). \tag{7.23}$$

A query data instance $x^{(q)}$ is then labeled as an anomaly if:

$$\log(p(x^{(q)})) < \varepsilon^*. \tag{7.24}$$

7.5 HASI Prediction via Deep Metric Learning Based Anomaly Detection (DML)

7.5.1 Network Structure

As discussed in Chapter 6, the model adopts a Siamese architecture similar to the one depicted in Figure 7.4. A key difference lies in the types of input and output. This model directly yields a binary decision: 1 indicates injury, while 0 denotes no injury prediction. Contrary to the traditional few-shot learning setting where the support set encompasses all potential classes, here the support set solely consists of non-injured screenings for the same player. In a 5-shot learning scenario, the support set comprises five sEMG data segments from the same athlete from instances where it is confirmed that the athlete did not sustain any injuries within the subsequent 8 days. Given that this is the concluding block of the proposed framework, the drill's class is predetermined, enabling automatic detection of the support set based on the drill type. For a 5-shot problem, at least five prior non-injured, or fit, measurements from the same drill involving the athlete are necessary.

Figure 7.4 demonstrates that the model employs the feature learning function F_θ to discern a representation space wherein vectors differentiate according to criteria established by the loss function. The model extracts features directly from the raw data; therefore, both query and support set values consist of segmented sEMG streams. The function F_θ is identical to the feature extractor referenced in Chapter 6. It yields feature vectors, which are then fed into the distance-based relation function G_ϕ. The function G_ϕ returns a binary injury prediction based on the distance of the query to the support set. This model is trained to distinguish features where data from fit athletes converge while anomalous data deviates. Relation learning identifies centroids separate for each drill class and subsequently applies the distance function, detailed in the subsequent subsection.

7.5.2 Training the Model: Loss Function Analysis

The distance between two vectors x_1 and x_2 is given by

$$\|x_1 - x_2\|_2^2 = \sum_{i=1}^{n} (x_1[i] - x_2[i])^2, \tag{7.25}$$

where n represents the vectors' dimensionality.

Unlike traditional triplet learning, this model adopts a distinctive approach to selecting triplets. The element x_i corresponds to post-injury sEMG data of a player. This post-injury data is exclusively associated with structural injuries as outlined in Chapter 2. The element x_s is linked to a safe MD+1 screening of a player, ensuring that no injury occurred within (3-8) days following the screening.

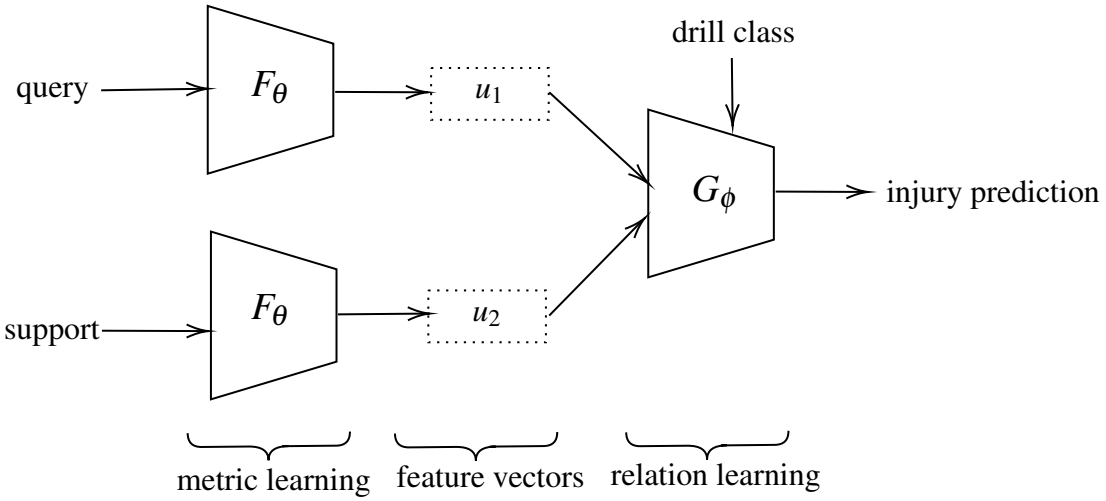


Fig. 7.4 Block diagram illustrating the deep metric learning framework for anomaly detection. Raw sEMG data serve as the model's input, processed by a twin feature extractor F_θ with shared weights. These feature vectors subsequently inform the distance function G_ϕ, which produces the injury prediction. Importantly, the support set is restricted to drill types consistent with the query, underscoring the significance of identifying the drill class.

Meanwhile, x_r is an MD+1 data sample from a player who sustained an indirect injury after the screening. The feature vectors for each triplet are determined as follows:

$$\begin{aligned} u_s &= F(x_s), \\ u_r &= F(x_r), \\ u_i &= F(x_i). \end{aligned} \tag{7.26}$$

A centroid matrix, denoted as $\mathcal{C}$, comprises randomly initialised representation vectors with 64 dimensions. Thus, $\mathcal{C} \in \mathbb{R}^{C\times 64}$ is both learnt and updated during training, containing optimal centroids for each drill type $c \in C$.

The distances are calculated as:

$$d_s = \frac{||\mathcal{C}[c]-u_s||_2^2}{||\mathcal{C}[c]-u_s||_2^2+||\mathcal{C}[c]-u_r||_2^2+||\mathcal{C}[c]-u_i||_2^2}, \tag{7.27}$$

$$d_r = \frac{||\mathcal{C}[c]-u_r||_2^2}{||\mathcal{C}[c]-u_s||_2^2+||\mathcal{C}[c]-u_r||_2^2+||\mathcal{C}[c]-u_i||_2^2}, \tag{7.28}$$

$$d_i = \frac{||\mathcal{C}[c]-u_i||_2^2}{||\mathcal{C}[c]-u_s||_2^2+||\mathcal{C}[c]-u_r||_2^2+||\mathcal{C}[c]-u_i||_2^2}. \tag{7.29}$$

Here, $\mathcal{C}[c]$ represents the centroid associated with the drill class $c \in C$. Additionally, d_s, d_r, and d_i denote the normalised distances of feature vectors to the learnt centroid for safe, risky, and injured data samples, respectively, all belonging to the same player's drill type c.

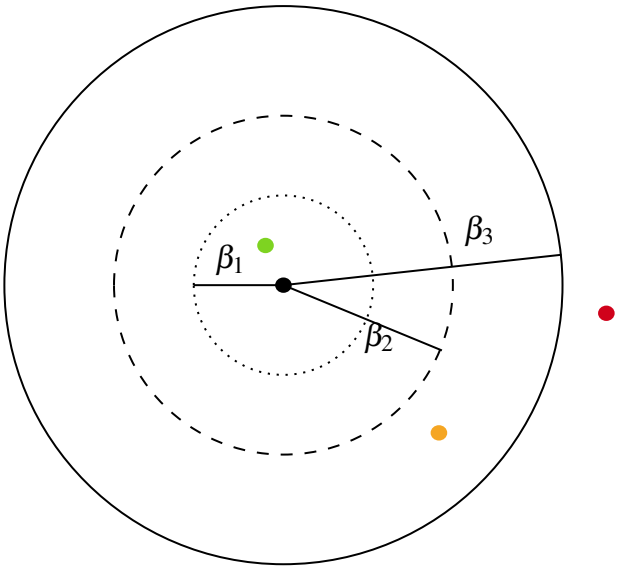


Fig. 7.5 A depiction of the centroid (symbolised by a black dot) with corresponding feature vectors for safe (u_s in green), risky (u_r in amber), and injured (u_i in red) states. The distance thresholds for safe (β_1), risky (β_2), and injured (β_3) are also highlighted. This diagram demonstrates how the loss function maintains safe vectors close to the centroid while setting the risky and injured vectors at regulated distances using their respective thresholds.

As depicted in Figure 7.5, data samples from safe versions of a player cluster around the centroid, within a nested sphere of radius β_1. This chapter posits that safe players exhibit similar feature vectors grouped around the centroid. When players' injury risk increases, variations in several intrinsic factors cause the feature vector to drift from the centroid, which β_2 regulates.

Contrary to regulated centroid triplet loss or proxy learning, only one centroid is learnt, representing the fit players. Fit players consistently exhibit similar patterns, whereas risky players lack a uniform anomaly pattern. Given the multifaceted ways players can distance themselves from their baseline, learning centroids for injured players is avoided. Instead, a sphere threshold β_2 is established further from β_1, with the condition $\beta_1 < \beta_2$.

As discussed in Section 7.2, sEMG recruitment patterns are expected to undergo changes following a structural injury. Consequently, players with structural injuries are anticipated to exhibit significant alterations in their sEMG recruitment patterns. These changes are likely to result in a greater deviation from the centroid. To enable model understanding of muscle behaviours when injured, post-injury athlete data is also incorporated with a sphere set further than risky data, given $\beta_3 > \beta_2$.

Taking all conditions into account, the model's loss is expressed as:

$$\begin{aligned} L_{\text{inj}} = &\max(d_s - \beta_1, 0) \\ &+ \max(\beta_2 - d_r, 0) \\ &+ \max(\beta_3 - d_i, 0), \end{aligned} \tag{7.30}$$

with set values of $\beta_1 = 0.1$, $\beta_2 = 0.5$, and $\beta_3 = 0.9$. These β values were chosen based on optimal performance on the dataset.

7.5.3 Injury Prediction through Model Inference

In this section, the proposed model's inference stage for injury prediction is outlined. A query is presented as a raw sEMG data segment from an MD+1 screening of an athlete. In parallel, a support set is defined, encompassing sEMG segments from the same drill undertaken by the aforementioned athlete. It is crucial to note that the support set consists of the athlete's uninjured screenings. Specifically, this indicates the athlete experienced no injuries up until the end of the subsequent MD.

In the initial phase of the inference process, the determination of $\mathcal{C}_{\text{selected}}$ is crucial. The centroid of the samples within the support set is defined as

$$\mathcal{C}_{\text{sup}} = \frac{\sum_{k=1}^{5} u_{f,k}}{5}, \tag{7.31}$$

considering the support set contains five samples. In this context, $\mathcal{C}_{\text{sup}}$ symbolises the centroid of the support set, while $u_{f,k}$ stands for the feature vectors of the support set, as signified by k. The subsequent step involves computing the distance between the support set centroid $\mathcal{C}_{\text{sup}}$ and the learnt centroid $\mathcal{C}[c]$ for a specific drill class c, given by

$$d_c = ||\mathcal{C}_{\text{sup}} - \mathcal{C}[c]||_2^2. \tag{7.32}$$

Conclusively, $\mathcal{C}_{\text{selected}}$ is determined based on the following criteria:

$$\mathcal{C}_{\text{selected}} = \begin{cases} \mathcal{C}[c] & \text{if } |d_c - \beta_2| \geq |d_c - \beta_1|, \\ \mathcal{C}_{\text{sup}} & \text{otherwise.} \end{cases} \tag{7.33}$$

The rationale behind the criterion in (7.33) is to tailor the model to each player's unique anatomical and physiological traits, significantly distinct from those of the general healthy player population. With only a five-sample support set, a player's centroid, $\mathcal{C}_{\text{sup}}$, may substantially deviate from the learned centroid, $\mathcal{C}[c]$, due to their unique muscle recruitment strategy. This deviation might position the player's centroid above the risk threshold, mistakenly identifying the player's normal data as anomalous, thus increasing the risk of false positives. To address this, the model must adapt by establishing a new centroid that reflects each player's individual traits, enhancing the system's accuracy and relevance in practical applications.

After determining the centroid, predictions are made based on the distance between the query and the centroid. Let $\mathcal{C}_{\text{selected}}$ be the chosen centroid and u_q be the query. The distance is then given by

$$d_q = \frac{||\mathcal{C}_{\text{selected}} - u_q||_2^2}{\sum_{s=1}^{5} ||u_{f,k} - \mathcal{C}_{\text{selected}}||_2^2 + ||\mathcal{C}_{\text{selected}} - u_q||_2^2}. \tag{7.34}$$

Subsequently, the distances corresponding to the safe and risky zones are calculated as $|d_q - \beta_1|$ and $|d_q - \beta_2|$, respectively. Given that a smaller distance signifies a greater likelihood of the respective

condition, these distances are negated as

$$d'_{\text{risky}} = -|d_q - \beta_2|, \tag{7.35}$$

$$d'_{\text{safe}} = -|d_q - \beta_1|. \tag{7.36}$$

The Softmax function can then be applied to these distances to convert them into probabilities. Specifically, the probability of the query being in the risky or safe zone is given by

$$P(\text{risky}) = \frac{e^{d'_{\text{risky}}}}{e^{d'_{\text{risky}}} + e^{d'_{\text{safe}}}}, \tag{7.37}$$

$$P(\text{safe}) = \frac{e^{d'_{\text{safe}}}}{e^{d'_{\text{risky}}} + e^{d'_{\text{safe}}}}, \tag{7.38}$$

where $P(\text{risky})$ and $P(\text{safe})$ denote the probabilities of the query being categorised into the risky and safe zones, respectively.

7.6 Experiments

This chapter details the experimental protocol for the proposed injury prediction methods referenced in Section 7.4 and Section 7.5. The dataset, denoted as D_f, comprises 4401 MD+1 measurements gathered over 2.5 seasons from 45 athletes across two EPL clubs. For clarity, MD+1 screenings without subsequent injuries up to the next MD are labelled as safe (0), whereas those leading to injuries within the same timeframe are labelled as risky (1). More information on the exercises, data collection methods, muscles assessed, and sensor placements can be found in Chapter 3. As observed in Table 7.2, the dataset is imbalanced: safe measurements account for 91.66% of the data, and risky measurements represent 8.33%, aligning with injury prevalence statistics mentioned in Chapter 1.

Table 7.2 Distribution of dataset size for safe and risky MD+1 measurements.

Total Measurements	**Safe (0)**	**Risky (1)**
4401	4034	367

It is important to acknowledge that although two clubs recorded data over most MD+1s during the 2.5 seasons, not every athlete was assessed on each MD+1. Only those who played for more than 60 minutes underwent measurements due to their exposure to gameplay load and injury risk. There were instances when MD+1 measurements could not be taken because of international away games, managerial decisions for days off, or unavailable staff. Consequently, this study does not cover all MD+1s for every player. It is equally noteworthy that injuries occurring during unmeasured MD+1s are not considered in this thesis. Over the 2.5 seasons, player departures and new arrivals meant the study captured a wide range of data from different athletes.

In this thesis, the HASI undergoes a collective analysis. Only indirect injuries, whether structural or functional, to muscles such as the soleus, hamstring, or adductor are categorised as risky (1) in this dataset. Certain drills specifically target particular muscles. For instance, the hamstring claw exercise pertains primarily to the hamstring and adductor, offering no insight into the soleus. Conversely, a bent knee soleus raise exclusively provides information about the soleus, excluding other muscles. Thus, labels for each drill are determined based on their target muscles. For illustration, if a player incurs a soleus injury following an MD+1 screening, their SL elevated glute bridge screening is not deemed risky but safe, as only exercises related to the soleus receive the risky classification. A comprehensive list detailing which exercises target specific muscle groups is available in Chapter 3.

Besides D_f, this chapter employs post-structural injury data from 20 athletes, sourced from the same clubs, aimed at return-to-play rehabilitation. These athletes underwent screenings every three days. Post-injury data often reveals pronounced anomalies, resulting from evident muscle tears. This contrasts with the risky data, where muscles have not sustained overt physical damage; generally, risky screenings exhibit better and more typical readings than post-injury data. This supplementary post-injury dataset aids the validation and training of the suggested methodologies. It is crucial to note that while the post-injury data enhances model performance, especially in deep learning, it is not employed for model testing. In this supplementary dataset, a total of 814 post-injury screenings are present.

In the dataset employed in this chapter, sEMG windows are approximately 8 seconds in duration, equivalent to 8192 samples, given that the sEMG sensors' sampling rate is 1 KHz. The 8-second duration for sEMG windows is chosen because most drills, including repetitions, typically last around this time. On occasions where players complete a drill in under 8 seconds, the window may contain baseline noise on either end. This noise does not influence the features extracted in both proposed methodologies. For sEMG data exceeding 8 seconds, it is trimmed to match the predetermined sampling rate of 8192 samples. For the Univariate Gaussian Anomaly Detection (UGAD) model, features are derived from these 8-second sEMG segments, while the deep learning model uses the entire raw sEMG segments as input. Consequently, a data instance in the dataset comprises 8129 samples for each muscle group, resulting in a shape of 8×8129, corresponding to channel size and sequence length, respectively.

In this chapter, no data augmentation techniques are applied. The Univariate Gaussian Anomaly Detection (UGAD) does not necessitate large data sets for distribution fitting and threshold selection. The Deep Metric Learning (DML) approach can expand its dataset since it utilises triplet combinations to discern the representation space. This allows for generating numerous triplets with varying combinations, enhancing the training dataset irrespective of the total number of dataset instances.

In the dataset D_f, each athlete-drill combination is considered a task, where each task comprises a set of data instances labeled as either safe or risky. As provided in Table 7.3, 6 exercises are selected from the D_f dataset, ensuring a sufficient number of MD+1 screenings to form more than 10 data samples for each task. This setup facilitates appropriate few-shot episodic testing, ensuring that every task can serve as both a support set and a query. To provide further context, we create 4196 triplets from each class, resulting in a total of 25176 triplets across 6 classes. Each triplet consists of 1 risky,

1 safe, and 1 injured screening of the same drill. Risky and safe measurements come from the same task, meaning the same player-drill combination; however, the injured data comes from the same drill only, as we may not have the injured data for each player. For each class, 16 distinct risky samples are chosen, with the exception of the Hamstring Claw class, where only 8 risky samples are selected due to relatively limited available data. In total, there are 220 tasks, with 88 tasks allocated for training and 132 tasks reserved for testing. The distribution of tasks is as follows: 34% for training, 60% for testing, and 6% for validation.

Table 7.3 Total number of players, and counts of risky and safe measurements per drill

Drill	Total Players	Risky	Safe
SL Elevated Glute Bridge	42	83	929
Isometric Prone Squeeze at 0°	43	83	865
60° Adductor Squeeze	42	38	615
SL Prone Curls	43	82	810
Hamstring Claw	14	47	215
SL Bent-Knee Soleus Raise	36	34	600

This method ensures a clear separation between the datasets used for training/validation and those used for testing, mirroring real-world scenarios where sports models must quickly adapt to new athlete-drill combinations without prior training. Since the model training and testing phases are realized by separating tasks, the model encounters new athlete-drill combinations in the training set, enhancing its adaptability to new tasks, meaning at least a new unseen subject or a new unseen drill. This replicates real-world scenarios where models need to generalize effectively. By conducting 10 iterations of this experiment, where tasks are randomly assigned to training or testing phases, we ensure robust model testing. After each iteration, we average the performance metrics to provide a comprehensive assessment of the model's capabilities. Consistency is maintained by keeping the test dataset constant across all iterations, facilitating reliable comparisons of model performance.

UGAD Training Methodology

The training for UGAD encompasses two primary phases. In the initial phase, Gaussian distributions for each feature are established based solely on normal (safe) data instances. Subsequently, in the second phase, a threshold is chosen using a validation set comprising both risky (anomaly) and safe (normal) data. To ensure a consistent comparison with the episodic testing approach, the support set used during the test phase is incorporated into the validation set.

DML Training Methodology

The training process for DML involves utilising triplets. As elucidated in Section 7.5, each triplet should comprise a sample from the risky label, a sample from the safe label, and a sample from the

post-injury label. Importantly, the data instances representing risky and safe within a triplet originate from the same task. However, the post-injury data is arbitrarily chosen from the post-injury dataset. This structure ensures that the distances between the safe and risky samples significantly influence representation learning F_θ. Thus, the model prioritises inter-task knowledge transfer, rather than generalising across the entire dataset, underscoring its meta-learning capability. The inclusion of post-injury data offers a reference for the model, guiding the direction in which to adjust the representation of risky data in more severe situations.

For the training, 4096 triplets are prepared. A validation set aids in averting overfitting during this phase. Contrasting with the UGAD training approach, this methodology neither incorporates the support set during the training nor the testing phase. However, the support set becomes instrumental during the model testing.

7.7 Results

Table 7.4 Test set performance metrics for the DML and UGAD models.

Metric	DML	UGAD
Accuracy	97.63%	79.37%
Recall	96.55%	77.78%
Precision	70.00%	60.87%
Specificity	97.69%	80.00%
NPV	99.80%	90.00%
FDR	30.00%	39.13%
FOR	0.20%	10.00%
F1 score	81.1594	68.2927

In the task of classifying athletes' conditions as either risky (1) or safe (0), two models are proposed: DML and UGAD. The results of their performance on the test set are presented in Table 7.4. Given the nature of the dataset, where the instances of injuries (labelled as 1) are considerably less frequent than non-injuries (labelled as 0), accuracy alone does not serve as a robust indicator of model performance. While both models exhibit satisfactory accuracy levels, the DML model particularly stands out with an impressive accuracy of 97.63%. However, due to the class imbalance, it is essential to consider other metrics that factor in both the positive and negative classes.

The F1 score, which is a harmonised measure of precision and recall, provides a comprehensive evaluation in such scenarios. The DML model achieves an F1 score of 81.1594, demonstrating its superior capability in estimating potential injuries. This high score is pivotal for injury prevention, given the repercussions of undetected risks.

Delving deeper into the critical errors, false positives (FP) and false negatives (FN) are of significant concern in this domain. An FP, termed as a false alarm, indicates a predicted injury risk when none is

present, potentially leading to unwarranted preventive measures. Conversely, FN is a critical oversight in this context. It signifies the model's failure to detect an impending injury, leading to missed precautionary measures. An undetected injury can have severe ramifications, including significant financial burdens for the club and a minimum two-week absence of the player, impacting team strategy and performance.

From the metrics presented, the DML model's recall of 96.55% highlights its ability to effectively identify a vast majority of actual injury risks, thus reducing FN errors. The precision of 70.00% points towards a certain propensity for FP errors, or false alarms. To further evaluate these false alarms, the FDR value, which stands at 30.00%, indicates that 30% of the positive identifications by the DML model might be false alarms. This reinforces the notion of a significant number of false alarms in predictions.

On the other side, considering missed detections, the FOR for the DML model is just 0.20%, meaning that only a small fraction of the negatives predicted by the model may actually be false. This low FOR complements the high recall, reinforcing the model's capability in minimising FN errors.

In contrast, the UGAD model has a recall of 77.78%, which, while respectable, is lower than that of the DML model. This suggests the UGAD model may not capture as many actual injury risks as the DML. Its precision stands at 60.87%, indicating a higher propensity for FP errors when compared to DML. The UGAD model's FDR is 39.13%, meaning there's an even greater chance of false alarms with its predictions than with DML. Additionally, the FOR for UGAD is at 10.00%, which is significantly higher than DML's, indicating a larger portion of its negative predictions might be incorrect.

In conclusion, while both models are proficient in estimating potential injuries, the DML model exhibits a notable edge, especially in injury prevention. Its higher F1 score and recall underscore its effectiveness. However, the elevated false alarms indicated by its precision warrants attention, suggesting potential areas for further refinement.

7.8 Discussion

In sports analytics, particularly in injury prediction, the implications of prediction errors are significant. When a false positive arises, which means predicting an injury when none exists, a player might be rested or excluded from a game for precaution. Although this may lead to short-term tactical disadvantages, it prevents more severe long-term repercussions.

On the other hand, a false negative, where an injury goes undetected, can have severe consequences. Allowing a vulnerable player to participate could result in actual injuries, sidelining them for a significant duration. Such outcomes not only affect the player's health but also impact the team's overall performance. Hence, from a safety and strategic perspective, tolerating occasional false positives is more acceptable than risking a false negative.

Several factors can lead to false positives. For instance, players identified as high-risk might receive limited game time, inherently reducing their injury exposure. Furthermore, while the model might identify a player as injury-prone in an upcoming match or within a short window (3-7 days), the actual

injury might occur after a more extended period, such as within two weeks. These intricacies introduce added layers to injury prediction, complicating the assessment.

Choosing an appropriate decision threshold is pivotal in classification tasks. When data corresponding to 'normal' and 'anomalies' (risky) are not distinctly separable, both Type I and Type II errors become inevitable. The location of this threshold can be strategically adjusted, based on the error type deemed more consequential. Given the substantial implications of false negatives in this domain, a preference might lean towards a more sensitive threshold. This approach may increase false positives but crucially reduces false negatives.

In summary, this chapter's model is designed with an intentional bias towards higher sensitivity. This approach prioritises player safety and aims to minimise the risks associated with undetected vulnerabilities.

7.9 Conclusion

In conclusion, the DML model attains an F1 score of 81.15 for HASI using solely the sEMG information. This result underscores the efficacy of the DML model and the proposed methodology in preventive measures for HASI. Additionally, testing the model on players with limited data points for specific player-drill (task) demonstrates the model's few-shot adaptability. This highlights how, through the episodic learning scheme, the model captures inter-task representation instead of merely generalising over the dataset.

The alternative proposed methodology, although not outperforming the DML, still achieves a commendable F1 score of 68.29. This supports the credibility and performance of the proposed signal processing-based feature selection. There remains potential for enhancing the signal processing-based features. Moreover, the UGAD model can be particularly effective and favoured in scenarios where injury or risky data is absent, requiring the model to infer potential injuries.

This chapter reveals that with the aid of sEMG data combined with deep learning methodologies, potential muscle injuries can be forecasted with significant accuracy. This underscores the importance and future potential of sEMG technology within the sports domain.

Chapter 8

Conclusion and Future Work

8.1 Summary of Contributions

This thesis presents a framework that exclusively utilises sEMG data from athletes to predict hamstring, adductor, and soleus indirect muscle injuries before their occurrence. Adopting a step-by-step approach, the thesis translates sEMG data into an injury prediction, encompassing everything from the initial recording to the probability phase. Each component of this framework contributes significantly to sports science and sEMG analysis. When integrated, these components achieve a predictive framework for HASI, attaining an F1 score of 81% on the D_f dataset. This work also aims to bridge the gap between academia and the sports industry, which often arises from feasibility issues and the unpredictable nature of real-life sports settings. Each component addresses a distinct challenge in utilising sEMG within the sports field. The thesis demonstrates that combining sEMG data and deep learning presents substantial potential in sports science, offering viable solutions to prevalent challenges.

Building on this foundation, the research addresses fundamental challenges in the real-life application of sEMG in sports. These challenges include drill-time labelling in uncontrolled sEMG data collection sessions, MIA contamination, and motion classification with limited data. The drill-time labelling model achieves 96.7% accuracy in detecting drill times from subwindows of 512 samples in sEMG recordings. Furthermore, the MIA cancellation method significantly enhances data quality, yielding a 17.20 dB improvement in SNR, which surpasses the top-performing state-of-the-art model by 7.01 dB. Additionally, the deep learning model developed for sEMG-based motion classification attains 98.80% accuracy on the publicly available benchmark dataset D_n, outperforming the best existing model by 8.66%. This model also exhibits high efficiency in the D_f dataset, achieving a 96.30% accuracy in classifying sports exercises. An advanced few-shot learning approach, utilising regulated centroid triplet loss, further enhances the model's accuracy to 90% in a 50-shot scenario, surpassing the best state-of-the-art model by 4%. This approach is useful for scenarios with limited datasets, a common constraint in the sports domain. Solutions to these obstacles, presented through deep learning techniques, are detailed in Chapters 4, 5, and 6. The proposed models are versatile, suitable for a range of applications beyond injury prevention, and address common challenges in sEMG analytics within the sports domain.

Moreover, the research introduces essential sEMG signal processing techniques in Chapter 7.3, providing crucial insights for evaluating athletes' risk and physiological states. The proposed metrics, including compensation, bilateral similarity, and explosiveness, offer simple and computationally efficient markers for monitoring physiological risks through sEMG data. These markers are especially practical in the sports domain, as they do not necessitate extensive datasets for model training, avoiding a significant limitation in this area. The proposed signal processing metrics have been independently validated with data collected from football clubs, confirming their practicality and efficacy.

8.2 Limitations

8.2.1 Absence of Inertial Measurement Unit (IMU) Sensors

The hardware utilized in this thesis is Neurocess sEMG, which does not include any additional IMU sensors such as those found in Delsys sEMG systems. IMU sensors provide 3-axis position data, offering an additional dimension of information regarding the kinematics of the athlete. This data could have been helpful in validating the drill-time labeling in Chapter 4. However, since the hardware used does not have IMU capabilities, the drill-time true labels were generated using the Neurocess app, where physiotherapists logged the onset and offset of drills. These true labels were subsequently validated through visual inspection of the raw data.

Furthermore, the IMU sensors can also be very useful in motion induced artefact (MIA) denoising in Chapter 5. MIA zones result from an impact on the muscle and sensors, which causes noise. With the use of IMU sensors, such artefact's onset times and magnitude can be easily detected. The addition of IMU sensor data could have enhanced the model performance or it could have been used for validation of MIA zone detection. Since the Neurocess sensors did not have any internal IMU sensors, we have labeled the MIA occurrence zones with expert view over the raw data.

8.2.2 Scope and Validation Limitations in Exercise Classification

The sEMG-based motion classification module described in Chapter 6 demonstrates high accuracy in distinguishing 9 distinct hamstring exercises, outperforming previous classification methods for hand gesture recognition tasks. However, the drills tested are very specific hamstring exercises, and the dataset has not been tested on other sports drills that are more dynamic, such as change of direction or sprinting exercises. This narrow scope is a significant limitation, as it may not fully represent the diversity of movements encountered in sports. Validation on other sets of less constrained tasks is crucial to establish the robustness and generalizability of the module for broader applications. Extending the validation to a larger number of sport-specific exercises would provide a more comprehensive assessment of the module's performance and potential in real-world sports scenarios.

8.2.3 Simplification of Denoising by Synthetic MIA Profile

The MIA model utilised in Chapter 5 presents two principal limitations: Firstly, it depends exclusively on a single scaling parameter b in (5.11) as the primary source of variability. This dependency may simplify the task for deep learning models, which need only to adapt to this singular source of variation. Secondly, the model assumes that the MIA is additive, neglecting potential multiplicative noise effects, such as those caused by positional shifts in electrodes relative to the sEMG source. These multiplicative effects can drastically compromise the signal's integrity, an issue not addressed by the proposed additive noise model.

8.2.4 Limitations in Injury Risk Labeling

The accurate classification of players' injury risk as either safe or risky presents significant challenges and inherent limitations. These difficulties primarily stem from the temporal discrepancies between risk identification and actual injury occurrence. For instance, a player may be correctly identified as being at high risk of injury yet not sustain an injury within the immediate observation window of 3-7 days. Consequently, this player might still sustain an injury in the subsequent two to three weeks. This situation illustrates a critical limitation: if no injury occurs within the designated short-term window, the player is classified as safe, potentially resulting in false positives. This classification challenge is compounded by the requirement to make timely decisions, typically aligned with the scheduling of upcoming matches within the next 3-7 days. The inability to extend this observation window adds further complexity to accurate risk assessment.

Moreover, there are additional factors influencing the true labeling of injury risk. Players identified as high risk might be allotted reduced playing time, inherently decreasing their exposure to injury-inducing scenarios. Furthermore, players perceiving themselves to be at high risk might consciously limit their exertion, thereby reducing their injury likelihood. These behaviors further complicate the predictive accuracy of injury models.

8.2.5 Impact of Player Transfers on Dataset Continuity

Player transfers significantly impact the continuity and regularity of data collection, presenting a notable challenge for the HASI prediction models. These models rely on accumulating a substantial dataset for each player to establish a baseline metric. When players are transferred between teams—either incoming or outgoing—the consistency of the data collection is disrupted. This results in insufficient data for accurately training HASI prediction models, thereby limiting their effectiveness.

8.3 Future Methodological Work

8.3.1 Utilising GAN for Noise Contamination

As discussed in Section 8.2.3, the MIA model introduced in Chapter 5 possesses inherent limitations due to its reliance on a single parameter and its exclusive focus on additive noise, disregarding the complexities of multiplicative noise effects. To address these shortcomings, the application of a Generative Adversarial Network (GAN) can be proposed. In this model, the discriminator will be trained to differentiate between sEMG signals contaminated with real and synthetically generated MIA, while the generator will aim to convincingly integrate MIA into sEMG windows. This approach not only reduces the dependency on a single parameter to introduce randomness but also expands the noise model to encompass multiplicative effects, such as those induced by shifts in electrode positions, which are not adequately represented by the existing MIA model in Chapter 5.

8.3.2 Motion Classification Test on Different Datasets

The few-shot learning model developed using regulated centroid triplet loss ($\mathcal{L}_{\text{RCTL}}$), as proposed in Chapter 6, has been initially tested exclusively on the D_f dataset. This model holds potential for various other few-shot learning tasks. Future testing phases will include applying the model to publicly available sEMG datasets such as D_n, as well as benchmark datasets commonly used in the literature for evaluating few-shot learning models. Additionally, the performance of $\mathcal{L}_{\text{RCTL}}$ will be compared with recent state-of-the-art methodologies using these publicly available datasets, allowing for further refinement and validation of the model's efficacy.

8.3.3 Advancing Injury Prevention with Synergistic Methodologies

Given the challenges identified with bipolar sEMG in assessing differences between synergistic muscle groups due to anatomical variations and other factors [110], future research will explore coherence-based methodologies for enhanced muscle connectivity analysis. Inspired by findings from [97, 19, 92], the future work will aim to integrate these coherence-based approaches to improve the precision of injury prevention strategies. Synergistic muscle groups, which coordinate to perform movements or bear loads, must be understood in unison rather than isolation. Coherence-based methods provide deep insights into the complex dynamics and interactions within these groups, offering potential for more customized and effective therapeutic and preventive measures in both clinical and athletic settings.

8.3.4 Enhancing Centroid Calculation: Adopting Deep Learning Over Arithmetic Mean

The current DML model for HASI prediction, detailed in Chapter 7, derives the centroid of a support set by computing the arithmetic mean of its vectors. However, this approach may not be optimal, particularly in the context of few-shot learning where the support set typically comprises only 5 to 10

samples. Additionally, given the high inter-session variance observed in the D_f dataset, primarily due to fluctuations in athletes' fatigue levels and physical load, the arithmetic mean might not accurately represent the true centroid of the data distribution when samples are limited. To address this, a deep learning-based approach could be employed for more effective centroid estimation in small sample scenarios. This method would involve leveraging a neural network to predict the centroid from the limited data available in the support set. Further enhancement could be achieved by incorporating a meta-learning strategy, which learns to predict centroids through episodic training. This training would involve various combinations of athlete-drill scenarios and their corresponding support sets, adapting the model to handle diverse and limited data more accurately.

8.3.5 Incorporating Multiple Drills Simultaneously into HASI Prediction Model

The current DML model for HASI prediction, as discussed in Chapter 7, utilises a singular drill from a recording session that comprises multiple drills, as detailed in Chapter 3. This model compares feature vectors from the chosen drill against a support set composed of similar drills performed by the same athlete to generate an injury risk score based on the distance to the healthy average of the support set. Future methodological advancements could involve extending the model to aggregate feature vectors from all the drills of the same recording session, assessing them against their respective support sets in a unified manner. The model is anticipated to yield a more robust prediction score by leveraging the collective insights gathered from multiple drills. This is based on the understanding that each drill encapsulates unique information about muscle engagement and potential risk factors. Thus, analysing the data not only for each drill individually but also between each drill in an integrated technique could offer a more comprehensive view of the athlete's condition. Future research will explore this comprehensive approach to enhance the model's predictive precision for HASI by integrating the entire range of exercises from a single session into the risk evaluation process.

8.3.6 Incorporating Temporal Variations into HASI Prediction Model

The framework introduced in this thesis offers HASI prediction using safe data (healthy versions of athletes) as the support set to distinguish between riskier versions through deep learning metrics. However, the model proposed in this research does not account for the temporal sequence of occurrences, specifically the evolution of a player's measurements over time. Understanding the direction in which an athlete's condition is moving—whether improving or deteriorating—could be crucial for predictions. This involves analysing the direction of the feature vector to enhance the model's predictive performance. In cases where an athlete exhibits a higher state of anomaly but shows a positive trajectory, indicating improvement, this would suggest a decreased level of risk. Incorporating this temporal dimension will provide a more nuanced understanding of an athlete's condition, thereby enhancing the accuracy and reliability of HASI risk predictions. Future work will focus on assessing such temporal variations and integrating them into the HASI prediction model.

8.4 Future Application Work

8.4.1 Broadening the Scope of HASI

This thesis's focus lies exclusively on predicting HASI. However, there is potential to explore injuries related to other muscles or diverse types of injuries and even different sports. Transferring the knowledge acquired from one sport to another can be advantageous, especially when data is limited. A significant avenue for future research, contingent upon gathering sufficient data, is to extend the model to encompass a wider array of injuries, diverse muscle groups, and various sports disciplines.

8.4.2 Tailoring Predictions for Female Athletes

The framework introduced in this thesis uses sEMG signals exclusively to forecast HASI muscle injuries among male athletes. This methodology does not yet cater to female athletes, a group with unique physiological considerations such as hormonal changes and menstrual cycle impacts. The direction for future research will be to integrate menstrual cycle data and hormonal information with sEMG signals. The objective will be to employ an appropriate methodology that can effectively predict injuries specific to women athletes by considering these combined data sources. This enhancement will focus on providing a more comprehensive and targeted approach to understanding and mitigating injury risks for female athletes.

8.4.3 Return-to-Play Decisions Following Rehab

While injury prevention remains paramount, the decision regarding an athlete's readiness to return to play is equally crucial. Inappropriate timings for return to play can lead to a recurrence of injuries. The same distance-based metric learning approach described in this thesis could be adapted for the opposite scenario. Instead of predicting an impending injury, the model could identify transitions from abnormal conditions back to normality, guiding return-to-play decisions more effectively.

References

[1] Abdoli Eramaki, M., Damecour, C., Christenson, J., and Stevenson, J. (2012). The Effect of Perspiration on the SEMG Amplitude and Power Spectrum. *Journal of Electromyography and Kinesiology*, 22(6):908–913.

[2] Agur, A. M., Ng-thow hing, V., Ball, K. A., Fiume, E., and Mckee, N. H. (2003). Documentation and Three-dimensional Modelling of Human Soleus Muscle Architecture. *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, 16(4):285–293.

[3] Allard, U. C., Nougarou, F., Fall, C. L., Giguère, P., Gosselin, C., Laviolette, F., and Gosselin, B. (2016). A Convolutional Neural Network for Robotic Arm Guidance Using SEMG Based Frequency-features. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2464–2470. IEEE.

[4] Amrutha, N. and Arul, V. (2017). A Review on Noises in EMG Signal and Its Removal. *Int. J. Sci. Res. Publ*, 7(5):23–27.

[5] Anthony G. Schache, Peter D. Blanch, Tim W. Dorn, Nicholas A. T. Brown, Doug Rosemond, and Marcus G. Pandy (2011). Effect of Running Speed on Lower Limb Joint Kinetics. *Medicine & Science in Sports & Exercise*, 43(7):1260–1271.

[6] Arnason, A., Sigurdsson, S. B., Gudmundsson, A., Holme, I., Engebretsen, L., and Bahr, R. (2004). Physical Fitness, Injuries, and Team Performance in Soccer. *Medicine & Science in Sports & Exercise*, 36(2):278–285.

[7] Arni Arnason, Stefan B. Sigurdsson, Arni Gudmundsson, Ingar Holme, Lars Engebretsen, and Roald Bahr (2004). Risk Factors for Injuries in Football. *The American Journal of Sports Medicine*, 32(1_suppl):5–16.

[8] Atzori, M., Cognolato, M., and Müller, H. (2016). Deep Learning with Convolutional Neural Networks Applied to Electromyography Data: A Resource for the Classification of Movements for Prosthetic Hands. *Frontiers in Neurorobotics*, 10:9.

[9] Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A.-g. M., Elsig, S., Giatsidis, G., Bassetto, F., and Müller, H. (2014a). Electromyography Data for Non-invasive Naturally-controlled Robotic Hand Prostheses. *Scientific Data*, 1(1):1–13.

[10] Atzori, M., Gijsberts, A., Kuzborskij, I., Elsig, S., Hager, A.-g. M., Deriaz, O., Castellini, C., Müller, H., and Caputo, B. (2014b). Characterization of a Benchmark Database for Myoelectric Movement Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(1):73–83.

[11] Barnes, C., Archer, D., Hogg, B., Bush, M., and Bradley, P. (2014). The Evolution of Physical and Technical Performance Parameters in the English Premier League. *International Journal of Sports Medicine*, pages 1095–1100.

[12] Basmajian, J. V. (1962). Muscles Alive. Their Functions Revealed by Electromyography. *Academic Medicine*, 37(8):802.

[13] Beiner, J. M. and Jokl, P. (2001). Muscle Contusion Injuries: Current Treatment Options. *Jaaos-journal of the American Academy of Orthopaedic Surgeons*, 9(4):227–237.

[14] Bengacemi, H., Abed meraim, K., Buttelli, O., Ouldali, A., and Mesloub, A. (2020). A New Detection Method for EMG Activity Monitoring. *Medical & Biological Engineering & Computing*, 58(2):319–334.

[15] Bengtsson, H., Ekstrand, J., Waldén, M., and Hägglund, M. (2013). Match Injury Rates in Professional Soccer Vary with Match Result, Match Venue, and Type of Competition. *The American Journal of Sports Medicine*, 41(7):1505–1510.

[16] Bennell, K., Wajswelner, H., Lew, P., Schall riaucour, A., Leslie, S., Plant, D., and Cirone, J. (1998). Isokinetic Strength Testing Does not Predict Hamstring Injury in Australian Rules Footballers. *British Journal of Sports Medicine*, 32(4):309–314.

[17] Blake, O. M. and Wakeling, J. M. (2015). Muscle Coordination Limits Efficiency and Power Output of Human Limb Movement under a Wide Range of Mechanical Demands. *Journal of Neurophysiology*, 114(6):3283–3295.

[18] Blandford, L., Theis, N., Charvet, I., and Mahaffey, R. (2018). Is Neuromuscular Inhibition Detectable in Elite Footballers during the Nordic Hamstring Exercise? *Clinical Biomechanics*, 58:39–43.

[19] Boonstra, T. W., Faes, L., Kerkman, J. N., and Marinazzo, D. (2019). Information Decomposition of Multichannel EMG to Map Functional Interactions in the Distributed Motor System. *Neuroimage*, 202:116093.

[20] Brady Green, Matthew N Bourne, Nicol Van Dyk, and Tania Pizzari (2020). Recalibrating the Risk of Hamstring Strain Injury (HSI): A 2020 Systematic Review and Meta-analysis of Risk Factors for Index and Recurrent Hamstring Strain Injury in Sport. *British Journal of Sports Medicine*, pages bjsports–2019–100983.

[21] Buhmann, R. L., Trajana, G., Kerr, G., and Shield, A. (2020). Voluntary Activation and Reflex Responses after Hamstring Strain Injury. *Medicine and Science in Sports and Exercise*, 52(9):1862–1869.

[22] Camilla L. Brockett, David L. Morgan, and Uwe Proske (2004). Predicting Hamstring Strain Injury in Elite Athletes. *Medicine & Science in Sports & Exercise*, 36(3):379–387.

[23] Castellini, C., Artemiadis, P., Wininger, M., Ajoudani, A., Alimusaj, M., Bicchi, A., Caputo, B., Craelius, W., Dosen, S., Englehart, K., and et al. (2014). Proceedings of the First Workshop on Peripheral Machine Interfaces: Going Beyond Traditional Surface Electromyography. *Frontiers in Neurorobotics*, 8.

[24] Chattopadhyay, R., Krishnan, N. C., and Panchanathan, S. (2011). Topology Preserving Domain Adaptation for Addressing Subject Based Variability in SEMG Signal. In *2011 Aaai Spring Symposium*, pages 4–9. AI Access Foundation.

[25] Chen, F., Wei, J., Xue, B., and Zhang, M. (2022). Feature Fusion and Kernel Selective in Inception-v4 Network. *Applied Soft Computing*, 119:108582.

[26] Chen, X., Li, Y., Hu, R., Zhang, X., and Chen, X. (2020). Hand Gesture Recognition Based on Surface Electromyography Using Convolutional Neural Network with Transfer Learning Method. *IEEE Journal of Biomedical and Health Informatics*, 25(4):1292–1304.

[27] Chiang, H.-t., Hsieh, Y.-y., Fu, S.-w., Hung, K.-h., Tsao, Y., and Chien, S.-y. (2019). Noise Reduction in ECG Signals Using Fully Convolutional Denoising Autoencoders. *IEEE Access*, 7:60806–60813.

[28] Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1251–1258.

[29] Clanton, T. O. and Coupe, K. J. (1998). Hamstring Strains in Athletes: Diagnosis and Treatment. *Jaaos-journal of the American Academy of Orthopaedic Surgeons*, 6(4):237–248.

[30] Clinic, M. (2023). Hamstring Injury: Symptoms and Causes. https://www.mayoclinic.org/diseases-conditions/hamstring-injury/symptoms-causes/syc-20372985. [Online; accessed 1-Sep-2023].

[31] Colberg, R. E., Aune, K. t., Choi, A. J., and Fleisig, G. s. (2015). Incidence and Prevalence of Musculoskeletal Conditions in Collegiate Tennis Athletes. *Medicine & Science in Tennis*, 20(3).

[32] Conceição, A., Silva, A. J., Barbosa, T., Karsai, I., and Louro, H. (2014). Neuromuscular Fatigue during 200 M Breaststroke. *Journal of Sports Science & Medicine*, 13(1):200.

[33] Costa, M., Goldberger, A. L., and Peng, C.-k. (2002). Multiscale Entropy Analysis of Complex Physiologic Time Series. *Physical Review Letters*, 89(6):068102.

[34] Côté allard, U., Fall, C. L., Drouin, A., Campeau lecours, A., Gosselin, C., Glette, K., Laviolette, F., and Gosselin, B. (2019). Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(4):760–771.

[35] Cram, J. R., Kasman, G. s., and Holtz, J. (1998). Introduction to Surface Electromyography.

[36] Cruz Ruiz, A., Pontonnier, C., Sorel, A., and Dumont, G. (2015). Identifying Representative Muscle Synergies in Overhead Football Throws. *Computer Methods in Biomechanics and Biomedical Engineering*, 18(sup1):1918–1919.

[37] D'aunno, D. s., Robinson, R. R., Smith, G. s., Thomason, D. B., and Booth, F. W. (1992). Intermittent Acceleration As a Countermeasure to Soleus Muscle Atrophy. *Journal of Applied Physiology*, 72(2):428–433.

[38] De Luca, C. J. (1997). The Use of Surface Electromyography in Biomechanics. *Journal of Applied Biomechanics*, 13(2):135–163.

[39] De Luca, C. J., Gilmore, L. D., Kuznetsov, M., and Roy, S. H. (2010). Filtering the Surface EMG Signal: Movement Artifact and Baseline Noise Contamination. *Journal of Biomechanics*, 43(8):1573–1579.

[40] Dellaserra, C. L., Gao, Y., and Ransdell, L. (2014). Use of Integrated Technology in Team Sports: A Review of Opportunities, Challenges, and Future Directions for Athletes. *The Journal of Strength & Conditioning Research*, 28(2):556–573.

[41] Delsys (2023). Trigno Sensors. https://delsys.com/trigno-avanti/#techspecs. [Online; accessed 27-November-2023].

[42] Di Nardo, F., Nocera, A., Cucchiarelli, A., Fioretti, S., and Morbidoni, C. (2022). Machine Learning for Detection of Muscular Activity from Surface EMG Signals. *Sensors*, 22(9):3393.

[43] Dieter, B. P., Mcgowan, C. P., Stoll, S. K., and Vella, C. a. (2014). Muscle Activation Patterns and Patellofemoral Pain in Cyclists. *Medicine & Science in Sports & Exercise*, 46(4):753–761.

[44] Ding, Z., Yang, C., Tian, Z., Yi, C., Fu, Y., and Jiang, F. (2018). SEMG-based Gesture Recognition with Convolution Neural Networks. *Sustainability*, 10(6):1865.

[45] D.l. Morgan (1990). New Insights into the Behavior of Muscle during Active Lengthening. *Biophysical Journal*, 57(2):209–221.

[46] Du, Y., Jin, W., Wei, W., Hu, Y., and Geng, W. (2017). Surface EMG-based Inter-session Gesture Recognition Enhanced by Deep Domain Adaptation. *Sensors*, 17(3):458.

[47] Eirale, C., Tol, J., Farooq, A., Smiley, F., and Chalabi, H. (2013). Low Injury Rate Strongly Correlates with Team Success in Qatari Professional Football. *British Journal of Sports Medicine*, 47(12):807–808.

[48] Ekstrand, J. (2013). Keeping Your Top Players on the Pitch: The Key to Football Medicine at a Professional Level.

[49] Ekstrand, J., Bengtsson, H., Waldén, M., Davison, M., Khan, K. M., and Hägglund, M. (2022). Hamstring Injury Rates Have Increased during Recent Seasons and now Constitute 24% of all Injuries in Men's Professional Football: The UEFA Elite Club Injury Study from 2001/02 to 2021/22. *British Journal of Sports Medicine*.

[50] Ekstrand, J., Hägglund, M., and Waldén, M. (2011a). Epidemiology of Muscle Injuries in Professional Football (Soccer). *The American Journal of Sports Medicine*, 39(6):1226–1232.

[51] Ekstrand, J., Hägglund, M., and Waldén, M. (2011b). Injury Incidence and Injury Patterns in Professional Football: The UEFA Injury Study. *British Journal of Sports Medicine*, 45(7):553–558.

[52] Eliakim, E., Morgulev, E., Lidor, R., and Meckel, Y. (2020). Estimation of Injury Costs: Financial Damage of English Premier League Teams' Underachievement due to Injuries. *BMJ Open Sport & Exercise Medicine*, 6(1):e000675.

[53] Enoka, R. M. (1994). *Neuromechanical Basis of Kinesiology*. Human kinetics Champaign, IL.

[54] Ergeneci, M., Bayram, E., Yarkin Yildiz, A., Carter, D., and Kosmas, P. (2023). sEMG Motion Classification Via Few-Shot Learning With Applications To Sports Science. *Authorea Preprints*.

[55] Ergeneci, M., Carter, D., and Kosmas, P. (2022). sEMG Onset Detection Via Bidirectional Recurrent Neural Networks with Applications to Sports Science. *IEEE Sensors Journal*, 22(19):18751–18761.

[56] Falese, L., Della Valle, P., and Federico, B. (2016). Epidemiology of Football (Soccer) Injuries in the 2012/2013 and 2013/2014 Seasons of the Italian Serie A. *Research in Sports Medicine*, 24(4):426–432.

[57] Farina, D., Fosci, M., and Merletti, R. (2002). Motor Unit Recruitment Strategies Investigated by Surface EMG Variables. *Journal of Applied Physiology*, 92(1):235–247.

[58] Farina, D., Jiang, N., Rehbaum, H., Holobar, A., Graimann, B., Dietl, H., and Aszmann, O. C. (2014). The Extraction of Neural Information from the Surface EMG for the Control of Upper-limb Prostheses: Emerging Avenues and Challenges. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4):797–809.

[59] Farina, D., Merletti, R., and Enoka, R. M. (2004). The Extraction of Neural Strategies from the Surface EMG. *Journal of Applied Physiology*, 96(4):1486–1495.

[60] Farrell, T. R. and Weir, R. F. (2007). The Optimal Controller Delay for Myoelectric Prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(1):111–118.

[61] Feldman, K., Franck, C., and Schauerte, C. (2018). Management of a Nonathlete with a Traumatic Groin Strain and Osteitis Pubis Using Manual Therapy and Therapeutic Exercise: A Case Report. *Physiotherapy Theory and Practice*.

[62] Fleming, N., Walters, J., Grounds, J., Fife, L., and Finch, a. (2015). Acute Response to Barefoot Running in Habitually Shod Males. *Human Movement Science*, 42:27–37.

[63] Freriks, B. and Hermens, H. (2000). *European Recommendations for Surface Electromyography: Results of the Seniam Project*. Roessingh Research and Development.

[64] Garrett Jr, W. E., Rich, F. R., Nikolaou, P. K., and Vogler 3rd, J. (1989). Computed Tomography of Hamstring Muscle Strains. *Medicine and Science in Sports and Exercise*, 21(5):506–514.

[65] Gary Henderson, Christopher A. Barnes, and Matthew D. Portas (2010). Factors Associated with Increased Propensity for Hamstring Injury in English Premier League Soccer Players. *Journal of Science and Medicine in Sport*, 13(4):397–402.

[66] Ge, W. (2018). Deep Metric Learning with Hierarchical Triplet Loss. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 269–285.

[67] Geng, W., Du, Y., Jin, W., Wei, W., Hu, Y., and Li, J. (2016). Gesture Recognition by Instantaneous Surface EMG Images. *Scientific Reports*, 6(1):36571.

[68] Gokcesu, K., Ergeneci, M., Ertan, E., Alkilani, A. Z., and Kosmas, P. (2018). An SEMG-based Method to Adaptively Reject the Effect of Contraction on Spectral Analysis for Fatigue Tracking. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers*, pages 80–87.

[69] Gordon, A. M., Huxley, A. F., and Julian, F. J. (1966). The Variation in Isometric Tension with Sarcomere Length in Vertebrate Muscle Fibres. *The Journal of Physiology*, 184(1):170–192.

[70] Green, B. and Pizzari, T. (2017). Calf Muscle Strain Injuries in Sport: A Systematic Review of Risk Factors for Injury. *British Journal of Sports Medicine*, 51(16):1189–1194.

[71] Gudelis, M., Pruna, R., Trujillano, J., Lundblad, M., and Khodaee, M. (2023). Epidemiology of Hamstring Injuries in 538 Cases from an FC Barcelona Multi Sports Club. *The Physician and Sportsmedicine*, pages 1–8.

[72] Hägglund, M., Waldén, M., and Ekstrand, J. (2005). Injury Incidence and Distribution in Elite Football—a Prospective Study of the Danish and the Swedish Top Divisions. *Scandinavian Journal of Medicine & Science in Sports*, 15(1):21–28.

[73] Hägglund, M., Waldén, M., and Ekstrand, J. (2006). Previous Injury As a Risk Factor for Injury in Elite Football: A Prospective Study over Two Consecutive Seasons. *British Journal of Sports Medicine*, 40(9):767–772.

[74] Hägglund, M., Waldén, M., and Ekstrand, J. (2013). Risk Factors for Lower Extremity Muscle Injury in Professional Soccer: The UEFA Injury Study. *The American Journal of Sports Medicine*, 41(2):327–335.

[75] Hawkins, R. D., Hulse, M., Wilkinson, C., Hodson, a., and Gibson, M. (2001). The Association Football Medical Research Programme: An Audit of Injuries in Professional Football. *British Journal of Sports Medicine*, 35(1):43–47.

[76] Hegyi, A., Goncalves, B., Finni Juutinen, T., and Cronin, N. (2019). Individual Region-and Muscle-specific Hamstring Activity at Different Running Speeds. *Medicine and Science in Sports and Exercise*, 51(11).

[77] Henneman, E., Somjen, G., and Carpenter, D. O. (1965). Functional Significance of Cell Size in Spinal Motoneurons. *Journal of Neurophysiology*, 28(3):560–580.

[78] Higashihara, A., Ono, T., Kubota, J., Okuwaki, T., and Fukubayashi, T. (2010). Functional Differences in the Activity of the Hamstring Muscles with Increasing Running Speed. *Journal of Sports Sciences*, 28(10):1085–1092.

[79] Hodges, P. W. and Bui, B. H. (1996). A Comparison of Computer-based Methods for the Determination of Onset of Muscle Contraction Using Electromyography. *Electroencephalography and Clinical Neurophysiology/electromyography and Motor Control*, 101(6):511–519.

[80] Hodgson, J. a., Finni, T., Lai, A. M., Edgerton, V. R., and Sinha, S. (2006). Influence of Structure on the Tissue Dynamics of the Human Soleus Muscle Observed in MRI Studies during Isometric Contractions. *Journal of Morphology*, 267(5):584–601.

[81] Hu, J., Shen, L., and Sun, G. (2018a). Squeeze-and-excitation Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7132–7141.

[82] Hu, Y., Wong, Y., Wei, W., Du, Y., Kankanhalli, M., and Geng, W. (2018b). A Novel Attention-based Hybrid CNN-RNN Architecture for SEMG-based Gesture Recognition. *Plos One*, 13(10):e0206049.

[83] Hug, F., Vogel, C., Tucker, K., Dorel, S., Deschamps, T., Le Carpentier, é., and Lacourpaille, L. (2019). Individuals Have Unique Muscle Activation Signatures As Revealed during Gait and Pedaling. *Journal of Applied Physiology*, 127(4):1165–1174.

[84] Hughes Iv, C., Hasselman, C. t., Best, T. M., Martinez, S., and Garrett Jr, W. E. (1995). Incomplete, Intrasubstance Strain Injuries of the Rectus Femoris Muscle. *The American Journal of Sports Medicine*, 23(4):500–506.

[85] Huygaerts, S., Cos, F., Cohen, D. D., Calleja gonzález, J., Guitart, M., Blazevich, A. J., and Alcaraz, P. E. (2020). Mechanisms of Hamstring Strain Injury: Interactions between Fatigue, Muscle Activation and Function. *Sports*, 8(5):65.

[86] Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., Webb, G. I., Idoumghar, L., Muller, P.-a., and Petitjean, F. (2020). Inceptiontime: Finding Alexnet for Time Series Classification. *Data Mining and Knowledge Discovery*, 34(6):1936–1962.

[87] Jan Ekstrand, Armin Spreco, Johann Windt, and Karim M. Khan (2020). Are Elite Soccer Teams' Preseason Training Sessions Associated with Fewer In-season Injuries? A 15-year Analysis from the Union of European Football Associations (UEFA) Elite Club Injury Study. *The American Journal of Sports Medicine*, 48(3):723–729.

[88] Jaramillo yánez, A., Benalcázar, M. E., and Mena maldonado, E. (2020). Real-time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review. *Sensors*, 20(9):2467.

[89] Järvinen, T. A., Järvinen, T. L., Kääriäinen, M., Kalimo, H., and Järvinen, M. (2005). Muscle Injuries: Biology and Treatment. *The American Journal of Sports Medicine*, 33(5):745–764.

[90] Jhang, Y.-s., Wang, S.-t., Sheu, M.-h., Wang, S.-h., and Lai, S.-c. (2022). Channel-wise Average Pooling and 1d Pixel-shuffle Denoising Autoencoder for Electrode Motion Artifact Removal in ECG. *Applied Sciences*, 12(14):6957.

[91] Jiang, N., Dosen, S., Muller, K., and Farina, D. (2012). Myoelectric Control of Artificial Limbs—Is There a Need to Change Focus? [In the Spotlight]. *IEEE Signal Processing Magazine*, 29(5):152–150.

[92] Jiménez grande, D., Atashzar, s. F., Martinez Valdes, E., and Falla, D. (2021). Muscle Network Topology Analysis for the Classification of Chronic Neck Pain Based on EMG Biomarkers Extracted during Walking. *Plos One*, 16(6):e0252657.

[93] Karlsson, S. and Gerdle, B. (2001). Mean Frequency and Signal Amplitude of the Surface EMG of the Quadriceps Muscles Increase with Increasing Torque—A Study Using the Continuous Wavelet Transform. *Journal of Electromyography and Kinesiology*, 11(2):131–140.

[94] Karnam, N. K., Dubey, S. R., Turlapaty, A. C., and Gokaraju, B. (2022). EMGhandnet: A Hybrid CNN and Bi-LSTM Architecture for Hand Activity Classification Using Surface EMG Signals. *Biocybernetics and Biomedical Engineering*, 42(1):325–340.

[95] Kaya, M. and Bilge, H. ş. (2019). Deep Metric Learning: A Survey. *Symmetry*, 11(9):1066.

[96] Kelcy Tobey and Jonathan Mike (2018). Single-leg Glute Bridge. *Strength & Conditioning Journal*, 40(2):110–114.

[97] Kerkman, J. N., Daffertshofer, A., Gollo, L. L., Breakspear, M., and Boonstra, T. W. (2018). Network Structure of the Human Musculoskeletal System Shapes Neural Interactions on Multiple Time Scales. *Science Advances*, 4(6):eaat0497.

[98] Khushaba, R. N. (2014). Correlation Analysis of Electromyogram Signals for Multiuser Myoelectric Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4):745–755.

[99] Kiel, J. and Kaiser, K. (2018). Adductor Strain.

[100] Kim, M., Moon, Y., Hunt, J., Mckenzie, K. a., Horin, A., Mcguire, M., Kim, K., Hargrove, L. J., and Jayaraman, A. (2021). A Novel Technique to Reject Artifact Components for Surface EMG Signals Recorded during Walking with Transcutaneous Spinal Cord Stimulation: A Pilot Study. *Frontiers in Human Neuroscience*, 15:660583.

[101] Koch, G., Zemel, R., and Salakhutdinov, R. (2015). Siamese Neural Networks for One-shot Image Recognition. In *ICML Deep Learning Workshop*, volume 2. Lille.

[102] Konrad, P. (2005). The ABC of EMG. *A Practical Introduction to Kinesiological Electromyography*, 1(2005):30–5.

[103] Lawrence, M. a., Leib, D. J., Ostrowski, S. J., and Carlson, L. a. (2017). Nonlinear Analysis of an Unstable Bench Press Bar Path and Muscle Activation. *The Journal of Strength & Conditioning Research*, 31(5):1206–1211.

[104] Lee, J., Mcmanus, D. D., Merchant, S., and Chon, K. H. (2011). Automatic Motion and Noise Artifact Detection in Holter ECG Data Using Empirical Mode Decomposition and Statistical Approaches. *IEEE Transactions on Biomedical Engineering*, 59(6):1499–1506.

[105] Linssen, W. H., Stegeman, D. F., Joosten, E. M., Notermans, S. L., Van't Hof, M. a., and Binkhorst, R. a. (1993). Variability and Interrelationships of Surface EMG Parameters during Local Muscle Fatigue. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, 16(8):849–856.

[106] Lowery, M. M., Vaughan, C. L., Nolan, P. J., and O'malley, M. J. (2000). Spectral Compression of the Electromyographic Signal due to Decreasing Muscle Fiber Conduction Velocity. *IEEE Transactions on Rehabilitation Engineering*, 8(3):353–361.

[107] Ltd, B. (2023). Biometrics: Surface EMG Sensor. https://www.biometricsltd.com/surface-emg-sensor.htm. [Online; accessed 27-November-2023].

[108] Luis Llurda almuzara, Noé Labata lezaun, Carlos López-de celis, Ramón Aiguadé aiguadé, Sergi Romaní sánchez, Jacobo Rodríguez sanz, César Fernández-de-las peñas, and Albert Pérez bellmunt (2021). Biceps Femoris Activation during Hamstring Strength Exercises: A Systematic Review. *International Journal of Environmental Research and Public Health*, 18(16):8733.

[109] Martin Hägglund, Markus Waldén, Henrik Magnusson, Karolina Kristenson, Håkan Bengtsson, and Jan Ekstrand (2013). Injuries Affect Team Performance Negatively in Professional Football: An 11-year Follow-up of the UEFA Champions League Injury Study. *British Journal of Sports Medicine*, 47(12):738–742.

[110] Martinez Valdes, E., Negro, F., Falla, D., De Nunzio, A. M., and Farina, D. (2018). Surface Electromyographic Amplitude Does not Identify Differences in Neural Drive to Synergistic Muscles. *Journal of Applied Physiology*, 124(4):1071–1079.

[111] Matthew Buckthorpe, Steve Wright, Stewart Bruce low, Gianni Nanni, Thomas Sturdy, Aleksander Stephan Gross, Laura Bowen, Bill Styles, Stefano Della Villa, Michael Davison, and Mo Gimpel (2018). Recommendations for Hamstring Injury Prevention in Elite Football: Translating Research into Practice. *British Journal of Sports Medicine*, 53(7):449–456.

[112] Merletti, R. and Farina, D. (2016). *Surface Electromyography: Physiology, Engineering, and Applications*. John Wiley & Sons.

[113] Merletti, R., Rainoldi, A., and Farina, D. (2001). Surface Electromyography for Noninvasive Characterization of Muscle. *Exercise and Sport Sciences Reviews*, 29(1):20–25.

[114] Mert Ergeneci, Kaan Gokcesu, Erhan Ertan, and Panagiotis Kosmas (2018). An Embedded, Eight Channel, Noise Canceling, Wireless, Wearable SEMG Data Acquisition System with Adaptive Muscle Contraction Detection. *IEEE Transactions on Biomedical Circuits and Systems*, 12(1):68–79.

[115] M.m. Lowery and M.j. O'malley (2003). Analysis and Simulation of Changes in EMG Amplitude during High-level Fatiguing Contractions. *IEEE Transactions on Biomedical Engineering*, 50(9):1052–1062.

[116] Mohan, D. D., Sankaran, N., Fedorishin, D., Setlur, S., and Govindaraju, V. (2020). Moving in the Right Direction: A Regularization for Deep Metric Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14591–14599.

[117] Mornieux, G., Gollhofer, a., and Stapelfeldt, B. (2010). Muscle Coordination While Pulling up during Cycling. *International Journal of Sports Medicine*, 31(12):843–846.

[118] Mosler, A. B., Weir, A., Serner, A., Agricola, R., Eirale, C., Farooq, A., Bakken, A., Thorborg, K., Whiteley, R. J., Hölmich, P., and Others (2018). Musculoskeletal Screening Tests and Bony Hip Morphology Cannot Identify Male Professional Soccer Players at Risk of Groin Injuries: A 2-year Prospective Cohort Study. *The American Journal of Sports Medicine*, 46(6):1294–1305.

[119] Mueller wohlfahrt, H.-w., Haensel, L., Mithoefer, K., Ekstrand, J., English, B., Mcnally, S., Orchard, J., Van Dijk, C. N., Kerkhoffs, G. M., Schamasch, P., and Others (2013). Terminology and Classification of Muscle Injuries in Sport: The Munich Consensus Statement. *British Journal of Sports Medicine*, 47(6):342–350.

[120] Naderi, A., Moen, M. H., and Degens, H. (2020). Is High Soleus Muscle Activity during the Stance Phase of the Running Cycle a Potential Risk Factor for the Development of Medial Tibial Stress Syndrome? AProspective Study. *Journal of Sports Sciences*, 38(20):2350–2358.

[121] Netter, F. H. (2014). *Atlas of Human Anatomy, Professional Edition E-book: Including Netter-reference. Com Access with Full Downloadable Image Bank*. Elsevier Health Sciences.

[122] Neurocess (2023). Neurocess Official Website. https://www.neurocess.co. [Online; accessed 27-November-2023].

[123] NHS (2023). Hamstring Injury. https://www.nhs.uk/conditions/hamstring-injury/. [Online; accessed 1-Sep-2023].

[124] Np, B. N., Hug, F., Guével, A., Colloud, F., Lardy, J., and Dorel, S. (2018). Changes in Motor Coordination Induced by Local Fatigue during a Sprint Cycling Task. *Medicine and Science in Sports and Exercise*, 50(7):1394–1404.

[125] Opar, D. a., Williams, M. D., and Shield, A. J. (2012). Hamstring Strain Injuries: Factors That Lead to Injury and Re-injury. *Sports Medicine*, 42:209–226.

[126] Opar, D. a., Williams, M. D., Timmins, R. G., Dear, N. M., and Shield, A. J. (2013a). Knee Flexor Strength and Bicep Femoris Electromyographical Activity Is Lower in Previously Strained Hamstrings. *Journal of Electromyography and Kinesiology*, 23(3):696–703.

[127] Opar, D. a., Williams, M. D., Timmins, R. G., Dear, N. M., and Shield, A. J. (2013b). Rate of Torque and Electromyographic Development during Anticipated Eccentric Contraction Is Lower in Previously Strained Hamstrings. *The American Journal of Sports Medicine*, 41(1):116–125.

[128] Orchard, J., Farhart, P., and Leopold, C. (2004). Lumbar Spine Region Pathology and Hamstring and Calf Injuries in Athletes: Is There a Connection? *British Journal of Sports Medicine*, 38(4):502–504.

[129] Ozaki, Y., Aoki, R., Kimura, T., Takashima, Y., and Yamada, T. (2016). Characterizing Muscular Activities Using Non-negative Matrix Factorization from EMG Channels for Driver Swings in Golf. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (embc)*, pages 892–895. IEEE.

[130] Park, K.-h. and Lee, S.-w. (2016). Movement Intention Decoding Based on Deep Learning for Multiuser Myoelectric Interfaces. In *2016 4th International Winter Conference on Brain-computer Interface (BCI)*, pages 1–2. IEEE.

[131] Pedret, C., Rodas, G., Balius, R., Capdevila, L., Bossy, M., Vernooij, R. W., and Alomar, X. (2015). Return to Play after Soleus Muscle Injuries. *Orthopaedic Journal of Sports Medicine*, 3(7):2325967115595802.

[132] Peter Stanton and Craig Purdam (1989). Hamstring Injuries in Sprinting—the Role of Eccentric Exercise. *Journal of Orthopaedic & Sports Physical Therapy*, 10(9):343–349.

[133] Presland, J. D., Timmins, R. G., Maniar, N., Tofari, P. J., Kidgell, D. J., Shield, A. J., Dickson, J., and Opar, D. a. (2021). Muscle Activity and Activation in Previously Strain-injured Lower Limbs: A Systematic Review. *Sports Medicine*, 51:2311–2327.

[134] Pulici, L., Certa, D., Zago, M., Volpi, P., and Esposito, F. (2023). Injury Burden in Professional European Football (Soccer): Systematic Review, Meta-analysis, and Economic Considerations. *Clinical Journal of Sport Medicine*, 33(4):450–457.

[135] Qizhu, S., Yining, S., Xiangfeng, D., and Zuchang, M. (2005). Onset Determination of Muscle Contraction in Surface Electromyography Signals Analysis. In *2005 IEEE International Conference on Information Acquisition*, pages 4–pp. IEEE.

[136] Rahimian, E., Zabihi, S., Asif, A., Farina, D., Atashzar, S. F., and Mohammadi, A. (2021). FS-HGR: Few-shot Learning for Hand Gesture Recognition Via Electromyography. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1004–1015.

[137] Rahimian, E., Zabihi, S., Atashzar, s. F., Asif, A., and Mohammadi, A. (2019). SEMG-based Hand Gesture Recognition Via Dilated Convolutional Neural Networks. In *2019 IEEE Global Conference on Signal and Information Processing (globalsip)*, pages 1–5. IEEE.

[138] Rahman, M. Z. U., Karthik, G., Fathima, S. Y., and Lay ekuakille, a. (2013). An Efficient Cardiac Signal Enhancement Using Time–frequency Realization of Leaky Adaptive Noise Cancelers for Remote Health Monitoring Systems. *Measurement*, 46(10):3815–3835.

[139] Rahman, M. Z. U., Shaik, R. A., and Reddy, D. R. K. (2011). Efficient and Simplified Adaptive Noise Cancelers for ECG Sensor Based Remote Health Monitoring. *IEEE Sensors Journal*, 12(3):566–573.

[140] Reaz, M. B. I., Hussain, M. S., and Mohd yasin, F. (2006). Techniques of EMG Signal Analysis: Detection, Processing, Classification and Applications. *Biological Procedures Online*, 8:11–35.

[141] Rebella, G. (2015). A Prospective Study of Injury Patterns in Collegiate Pole Vaulters. *The American Journal of Sports Medicine*, 43(4):808–815.

[142] Roland, T., Baumgartner, W., Amsuess, S., and Russold, M. F. (2017). Capacitively Coupled EMG Detection via Ultra-low-power Microcontroller STFT. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 410–413.

[143] Ross A. Clark (2008). Hamstring Injuries: Risk Assessment and Injury Prevention. *Annals of the Academy of Medicine, Singapore*, 37 4:341–6.

[144] Sato, K., Nimura, A., Yamaguchi, K., and Akita, K. (2012). Anatomical Study of the Proximal Origin of Hamstring Muscles. *Journal of Orthopaedic Science*, 17:614–618.

[145] Schache, A. G., Dorn, T. W., Williams, G. P., Brown, N. A., and Pandy, M. G. (2014). Lower-limb Muscular Strategies for Increasing Running Speed. *Journal of Orthopaedic & Sports Physical Therapy*, 44(10):813–824. PMID: 25103134.

[146] Schafer, R. W. (2011). What Is a Savitzky-golay Filter? *IEEE Signal Processing Magazine*, 28(4):111–117.

[147] Seidel, E., Rother, M., Hartmann, J., Rother, I., Schaaf, T., Winzer, M., Fischer, A., and Regenspurger, K. (2012). Eccentric Exercise and Delayed Onset of Muscle Soreness (DOMS)–An Overview. *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin*, 22(02):57–63.

[148] Serner, A., Mosler, A. B., Tol, J. L., Bahr, R., and Weir, A. (2019). Mechanisms of Acute Adductor Longus Injuries in Male Football Players: A Systematic Visual Video Analysis. *British Journal of Sports Medicine*, 53(3):158–164.

[149] Silva, L., Vaz, J. R., Castro, M. A., Serranho, P., Cabri, J., and Pezarat correia, P. (2015). Recurrence Quantification Analysis and Support Vector Machines for Golf Handicap and Low Back Pain EMG Classification. *Journal of Electromyography and Kinesiology*, 25(4):637–647.

[150] Silvia Conforto, Tommaso D'alessio, and Stefano Pignatelli (1999). Optimal Rejection of Movement Artefacts from Myoelectric Signals by Means of a Wavelet Filtering Procedure. *Journal of Electromyography and Kinesiology*, 9(1):47–57.

[151] Simão, M., Mendes, N., Gibaru, O., and Neto, P. (2019). A Review on Electromyography Decoding and Pattern Recognition for Human-machine Interaction. *IEEE Access*, 7:39564–39582.

[152] Snell, J., Swersky, K., and Zemel, R. (2017). Prototypical Networks for Few-shot Learning. *Advances in Neural Information Processing Systems*, 30.

[153] Sole, G., Milosavljevic, S., Nicholson, H., and Sullivan, s. J. (2012). Altered Muscle Activation Following Hamstring Injuries. *British Journal of Sports Medicine*, 46(2):118–123.

[154] Steinhardt, C. R., Bettthauser, J., Hunt, C., and Thakor, N. (2018). Registration of EMG Electrodes to Reduce Classification Errors due to Electrode Shift. In *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4.

[155] Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P. H., and Hospedales, T. M. (2018). Learning to Compare: Relation Network for Few-shot Learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1199–1208.

[156] Taborri, J., Keogh, J., Kos, A., Santuz, A., Umek, A., Urbanczyk, C., Van Der Kruk, E., Rossi, S., and Others (2020). Sport Biomechanics Applications Using Inertial, Force, and EMG Sensors: A Literature Overview. *Applied Bionics and Biomechanics*, 2020.

[157] UEFA (2023). Injury Reasons: UEFA Medical Document. https://www.uefa.com/MultimediaFiles/Download/uefaorg/Medical/02/61/67/86/2616786_DOWNLOAD.pdf. [Online; accessed 1-Sep-2023].

[158] Urho M. Kujala, Sakari Orava, and Markku Järvinen (1997). Hamstring Injuries. *Sports Medicine*, 23(6):397–404.

[159] Venkatesan, C., Karthigaikumar, P., and Varatharajan, R. (2019). Fpga Implementation of Modified Error Normalized Lms Adaptive Filter for ECG Noise Removal. *Cluster Computing*, 22:12233–12241.

[160] Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., and Others (2016). Matching Networks for One Shot Learning. *Advances in Neural Information Processing Systems*, 29.

[161] Visconti, P., Gaetani, F., Zappatore, G., and Primiceri, P. (2018). Technical Features and Functionalities of Myo Armband: An Overview on Related Literature and Advanced Applications of Myoelectric Armbands Mainly Focused on Arm Prostheses. *International Journal on Smart Sensing and Intelligent Systems*, 11:1–25.

[162] Volpi, P. (2008). Epidemiology and Risk Factors in Soccer. *Medicina Dello Sport*, 61(2):111–119.

[163] Waldén, M., Hägglund, M., Orchard, J., Kristenson, K., and Ekstrand, J. (2013). Regional Differences in Injury Incidence in European Professional Football. *Scandinavian Journal of Medicine & Science in Sports*, 23(4):424–430.

[164] Wang, J., Tang, L., and Bronlund, J. E. (2019). Pattern Recognition-based Real Time Myoelectric System for Robotic Hand Control. In *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 1598–1605.

[165] Wang, Y., Yao, Q., Kwok, J. t., and Ni, L. M. (2020). Generalizing from a Few Examples: A Survey on Few-shot Learning. *ACM Computing Surveys (CSUR)*, 53(3):1–34.

[166] Wei, W., Dai, Q., Wong, Y., Hu, Y., Kankanhalli, M., and Geng, W. (2019a). Surface-electromyography-based Gesture Recognition by Multi-view Deep Learning. *IEEE Transactions on Biomedical Engineering*, 66(10):2964–2973.

[167] Wei, W., Wong, Y., Du, Y., Hu, Y., Kankanhalli, M., and Geng, W. (2019b). A Multi-stream Convolutional Neural Network for SEMG-based Gesture Recognition in Muscle-computer Interface. *Pattern Recognition Letters*, 119:131–138.

[168] Welford, B. (1962). Note on a Method for Calculating Corrected Sums of Squares and Products. *Technometrics*, 4(3):419–420.

[169] William E. Garrett (1990). Muscle Strain Injuries. *Medicine & Science in Sports & Exercise*, 22(4):436–443.

[170] Windt, J., Ekstrand, J., Khan, K. M., Mccall, A., and Zumbo, B. D. (2018). Does Player Unavailability Affect Football Teams' Match Physical Outputs? ATwo-season Study of the UEFA Champions League. *Journal of Science and Medicine in Sport*, 21(5):525–532.

[171] Winter, D. A. (2009). *Biomechanics and Motor Control of Human Movement*. John Wiley & Sons.

[172] Witvrouw, E., Danneels, L., Asselman, P., D'have, T., and Cambier, D. (2003). Muscle Flexibility As a Risk Factor for Developing Muscle Injuries in Male Professional Soccer Players: A Prospective Study. *The American Journal of Sports Medicine*, 31(1):41–46.

[173] Woo, S., Park, J., Lee, J.-y., and Kweon, i. s. (2018). Cbam: Convolutional Block Attention Module. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 3–19.

[174] Woods, C. (2004). The Football Association Medical Research Programme: An Audit of Injuries in Professional Football–Analysis of Hamstring Injuries. *British Journal of Sports Medicine*, 38(1):36–41.

[175] Xiong, D., Zhang, D., Zhao, X., and Zhao, Y. (2021). Deep Learning for EMG-based Human-machine Interaction: A Review. *IEEE/CAA Journal of Automatica Sinica*, 8(3):512–533.

[176] Xiong, P., Wang, H., Liu, M., and Liu, X. (2015). Denoising Autoencoder for Eletrocardiogram Signal Enhancement. *Journal of Medical Imaging and Health Informatics*, 5(8):1804–1810.

[177] Xu, Q., Quan, Y., Yang, L., and He, J. (2012). An Adaptive Algorithm for the Determination of the Onset and Offset of Muscle Contraction by EMG Signal Processing. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(1):65–73.

[178] Yadav, S. K., Sinha, R., and Bora, P. K. (2015). Electrocardiogram Signal Denoising Using Non-local Wavelet Transform Domain Filtering. *IET Signal Processing*, 9(1):88–96.

[179] Yaghoubi, M., Lark, S. D., Page, W. H., Fink, P. W., and Shultz, S. P. (2019). Lower Extremity Muscle Function of Front Row Rugby Union Scrummaging. *Sports Biomechanics*, 18(6):636–648.

[180] Yao, W., Fuglevand, R. J., and Enoka, R. M. (2000). Motor-unit Synchronization Increases EMG Amplitude and Decreases Force Steadiness of Simulated Contractions. *Journal of Neurophysiology*, 83(1):441–452.

[181] Young, A. J., Hargrove, L. J., and Kuiken, T. A. (2011). The Effects of Electrode Size and Orientation on the Sensitivity of Myoelectric Pattern Recognition Systems to Electrode Shift. *IEEE Transactions on Biomedical Engineering*, 58(9):2537–2544.

[182] Young, A. J., Hargrove, L. J., and Kuiken, T. A. (2012). Improving Myoelectric Pattern Recognition Robustness to Electrode Shift by Changing Interelectrode Distance and Electrode Configuration. *IEEE Transactions on Biomedical Engineering*, 59(3):645–652.

[183] Zabihi, S., Rahimian, E., Asif, A., and Mohammadi, A. (2022). TRAHGR: Transformer for Hand Gesture Recognition Via Electromyography. *Arxiv Preprint Arxiv:2203.16336*.

[184] Zagoruyko, S. and Komodakis, N. (2016). Wide Residual Networks. *Arxiv Preprint Arxiv:1605.07146*.

[185] Zhai, X., Jelfs, B., Chan, R. H., and Tin, C. (2017). Self-recalibrating Surface EMG Pattern Recognition for Neuroprosthesis Control Based on Convolutional Neural Network. *Frontiers in Neuroscience*, 11:379.

[186] Zhang, X. and Zhou, P. (2012). Sample Entropy Analysis of Surface EMG for Improved Muscle Activity Onset Detection against Spurious Background Spikes. *Journal of Electromyography and Kinesiology*, 22(6):901–907.

[187] Zia Ur Rehman, M., Waris, A., Gilani, S. O., Jochumsen, M., Niazi, I. K., Jamil, M., Farina, D., and Kamavuako, E. N. (2018). Multiday EMG-based Classification of Hand Motions with Deep Learning Techniques. *Sensors*, 18(8):2497.

[188] Zory, R., Molinari, F., Knaflitz, M., Schena, F., and Rouard, A. (2011). Muscle Fatigue during Cross Country Sprint Assessed by Activation Patterns and Electromyographic Signals Time–frequency Analysis. *Scandinavian Journal of Medicine & Science in Sports*, 21(6):783–790.

Appendix A

Details of sEMG Data Acquisition Systems

As detailed in Chapter 3, two datasets are utilized in this thesis: D_n, a publicly available dataset collected using the Delsys sEMG system, and D_f, collected specifically for this study using the Neurocess sEMG system.

A.1 Neurocess sEMG System Specifications

The Neurocess system, used for collecting the D_f dataset, comprises eight dry, active surface electromyography sensors. These sensors adhere to SENIAM and ISEK requirements and have the following specifications:

- Circular electrodes with a diameter of 10 mm are used.
- The inter-electrode distance is 20 mm center-to-center.
- The electrodes are made of a copper core and are gold plated.
- The device supports wireless data transmission via WiFi, with both UDP and TCP/IP protocols.
- The system has high analog to digital conversion performance with 16-bit resolution, 93 dB SNR, and -108 dB total harmonic distortion.
- It features separated digital and analog power planes for low noise and high effective number of bits (ENOB).
- The instrumentation amplifier has a common mode rejection ratio (CMRR) of 135 dB.
- The bandpass filter ranges from 20 to 450 Hz with a combination of 2-stage high-pass and 2-stage low-pass filters.
- The system has high input impedance.
- The dimensions of the device are 31 mm x 31 mm x 13 mm.

- The device includes adjustable amplification.
- It is equipped with an ARM Cortex M4 based microprocessor.
- The battery life of the device is up to 3 hours.

In Figure A.1, Neurocess sensors attached to medical-grade tape are shown. Figure A.2 depicts the application of Neurocess sensors on a subject. Finally, Figure A.3 visualizes the Neurocess systems in use on a training ground."

A.2 Comparison with Other sEMG Systems

In addition to Neurocess, several other solutions for sEMG data acquisition exist in the academic literature and commercial market, such as Biometrics (SX230), Delsys Trigno Lab, and Myo Armband. The basic hardware specifications of these systems are compared in Table A.1.

Table A.1 Comparison of the sEMG data acquisition systems

Systems	DT	Electrode Type/Material	Ch.	Gain	BW (Hz)	F_s (Hz)	Res. (bits)
Neurocess [122]	WiFi	Dry / Gold Plated Copper	8	500-3000	20-450	1000	16
Delsys [41]	RF	Dry / Silver	16	-	20-450 Hz	$\leq$4370	16
SX230 [107]	BLE	Dry / Stainless Steel	8	1000	20-460	1000	14
Myo Armband [161]	BLE	Dry / Stainless Steel	8	-	-	200	8

Fig. A.1 Illustration of Neurocess sensors on medical-grade adhesive tape prior to attachment to the body.

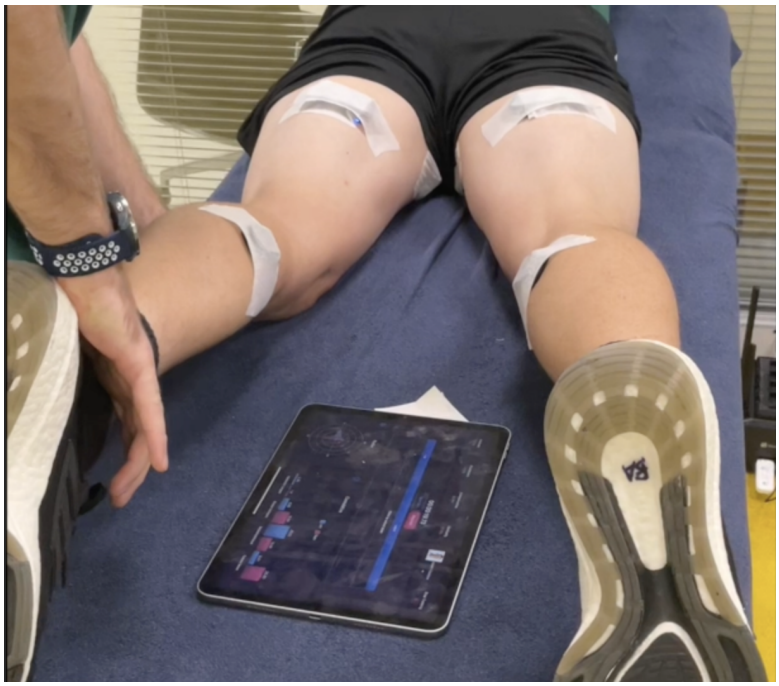


Fig. A.2 Illustration demonstrating the attachment of Neurocess sensors on the body. This image is solely for the purpose of showing sensor application and does not represent actual data collection for the D_f dataset.

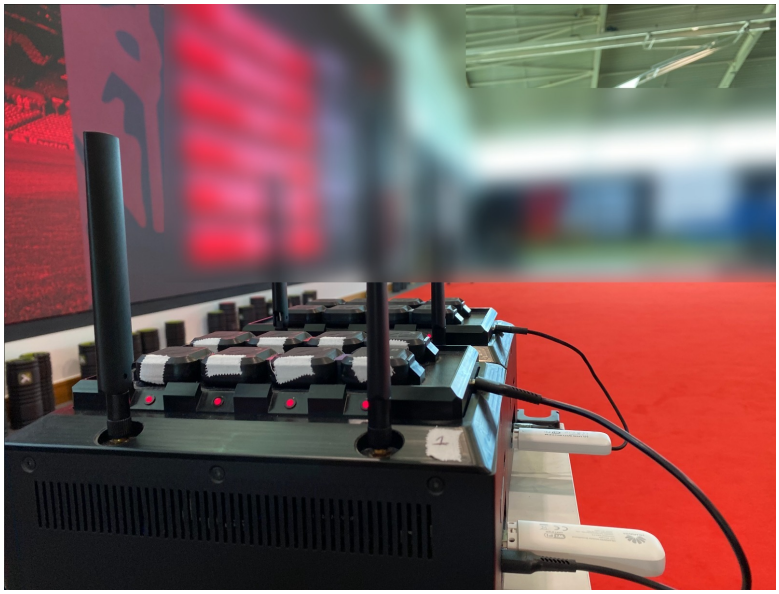


Fig. A.3 Illustration showing the Neurocess sensor system on the training ground.