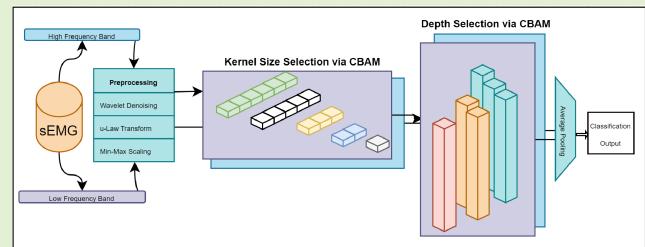


# Attention-Enhanced Frequency-Split Convolution Block for sEMG Motion Classification: Experiments on Premier League and Ninapro Datasets

Mert Ergeneci<sup>ID</sup>, Erkan Bayram<sup>ID</sup>, David Binningsley, Daryl Carter, and Panagiotis Kosmas<sup>ID</sup>, Senior Member, IEEE

**Abstract**—This article presents convolutional octave-band zooming-in with depth-kernel attention learning (COZDAL), a versatile deep learning model designed for surface electromyography (sEMG) motion classification. Specifically focusing on sports movements involving the hamstring muscle, the model employs attention mechanisms across various frequency bands, kernel sizes, and hidden layer depths. The proposed method has been extensively evaluated on the benchmark Ninapro dataset and a custom soccer dataset. The results demonstrate substantial improvements over the existing state-of-the-art models, with an accuracy of 95.30% on Ninapro DB2, outperforming the previous best by 3.29%, and an accuracy of 98.80% on Ninapro DB2-B, an 8.66% enhancement. Remarkably, COZDAL exhibits a performance accuracy of 96.30% on a soccer dataset gathered from 45 elite-level athletes representing two clubs in the English Premier League (EPL). This result, achieved without parameter tuning, highlights the model's adaptability and exceptional efficacy across diverse motion scenarios, sensors, subjects, and muscle types.



**Index Terms**—Attention, convolutional block attention module (CBAM), convolutional neural network (CNN), gesture recognition, motion classification, neural networks, sports science, surface electromyography (sEMG).

## I. INTRODUCTION

SURFACE electromyography (sEMG) captures electrical signals generated by muscle contraction, yielding critical

Manuscript received 27 November 2023; accepted 18 December 2023. Date of publication 28 December 2023; date of current version 13 February 2024. The associate editor coordinating the review of this article and approving it for publication was Dr. Zhenghua Chen. (Corresponding author: Mert Ergeneci.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the King's College London College Research Ethics Committee with reference number MRSP-20/21-20996.

Mert Ergeneci is with the Department of Engineering, Faculty of Natural, Mathematical, and Engineering Sciences, King's College London, WC2R 2LS London, U.K., and also with Neurocess Ltd., WC2A 2JR London, U.K. (e-mail: mert@neurocess.co).

Erkan Bayram is with the Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois Urbana-Champaign, Urbana, IL 61801 USA, and also with Neurocess Ltd., WC2A 2JR London, U.K. (e-mail: erkan@neurocess.co).

David Binningsley is with the Manchester United Football Club, M16 0RA Manchester, U.K. (e-mail: david.binningsley@manutd.co.uk).

Daryl Carter is with the Leeds United Football Club, LS11 0ES Leeds, U.K. (e-mail: daryl.carter@leedsunited.com).

Panagiotis Kosmas is with the Department of Engineering, Faculty of Natural, Mathematical, and Engineering Sciences, King's College London, WC2R 2LS London, U.K. (e-mail: panagiotis.kosmas@kcl.ac.uk).

Digital Object Identifier 10.1109/JSEN.2023.3345731

physiological and biomechanical insights [1], [2], [3], [4], [5], [6], [7]. Although widely used for hand gesture recognition (HGR) in prosthesis control, its potential in sports remains underexplored [8], [9], [10], [11]. Muscle dynamics vary with different movements, so motion classification using sEMG signals becomes indispensable in sports settings. Most studies of sEMG-based HGR utilize Ninapro, a publicly accessible benchmark dataset for HGR [6]. In this study, we extend the applications of sEMG beyond the Ninapro dataset, specifically focusing on sports movements involving the hamstring muscle. This focus is motivated by the high prevalence of hamstring strain injuries (HSIs) in athletes, especially soccer players, leading to significant time off the field [12], [13], [14], [15], [16], [17], [18]. We aim to enhance sEMG-based HSI evaluation by classifying hamstring-targeted movements, a step forward in risk assessment and injury prevention. We develop a deep learning model using sEMG data collected from hamstring exercises of first-team players from two English Premier League (EPL) clubs.

The task of sEMG-based motion classification can be achieved using feature engineering or feature learning, both relying on feature vector representations of the sEMG signal [6]. While feature engineering uses signal processing

algorithms, feature learning employs deep neural networks [19], [20]. Recent evidence from HGR studies suggests that feature learning outperforms feature engineering-based models in sEMG-based motion classification [21], [22], [23], [24], [25], [26], [27].

In feature learning for sEMG data, convolutional neural networks (CNNs) are preferred for distinguishing discriminative feature representations [19]. Studies, such as [21], demonstrate the application of CNNs for the classification of various hand gestures, outperforming traditional feature engineering methods utilizing support vector machines (SVMs). A common approach involves the integration of sEMG raw data streams into a matrix, which then serves as an input to a 2-D CNN for feature learning [22], [24], [28], [29]. Other works focus on using the spectrogram of sEMG, computed by the short-time Fourier transform (STFT), as a 2-D CNN input [28], [30]. This representation encapsulates both frequency- and time-domain information of the sEMG signal in a 2-D image form. A noteworthy comparison between the continuous wavelet transform (CWT) and the STFT reveals that the CWT performs better when used with a CNN [31]. In addition, the study explains the potential of slow-fusion CNNs in enhancing feature learning performance by establishing temporal dependencies between CWT images. In the pursuit of further improvement, [32] introduces a hybrid architecture combining CNN and recurrent neural network (RNN) wherein the sEMG data is temporally segmented and processed through parallel 2-D CNNs, with the output being directed into an RNN structure equipped with an attention mechanism. A subsequent model, EMGHandNet, combines 1-D CNN and BiLSTM-RNN structure, exhibiting improved results over Hu et al.'s model [32] and highlighting the importance of 1-D CNN in processing sEMG data [33]. In a further refinement to the 1-D CNN approach for sEMG HGR, Rahimian et al. [34] implemented a dilated 1-D CNN block, demonstrating superior performance over many existing models on the Ninapro DB2-B dataset. Recently, a novel transformer-encoder structure with multihead attention (MHA), TraHGR-Huge, has been introduced to improve HGR [35]. This model employs two parallel transformer structures to process normal and transposed input, enabling the learning of channelwise relations.

In motion classification, researchers often evaluate their model's performance using publicly available datasets such as Ninapro. However, no research currently assesses the model's performance using both Ninapro and a completely distinct dataset that includes various subject types, different sensor hardware (Neurocess), diverse muscle groups (lower extremity muscles), and motion classes. This article comprehensively evaluates the proposed model by testing its accuracy on the benchmark Ninapro dataset and examining its performance on a custom soccer dataset collected from professional athletes in the EPL. By doing so, we contribute valuable insights into motion classification and showcase our model's adaptability and generalization capabilities across different datasets and motion scenarios.

In recent studies, using the MHA mechanism has been shown to improve the performance of the HGR model [27], [35]. MHA applies attention in a parallel manner and then

merges the outcomes [36]. However, the time series and multichannel nature of sEMG data suggest that a convolutional block attention module (CBAM) is better suited to focus on informative features in both temporal and channel dimensions [37]. This is because time series data frequently exhibit local temporal patterns. CBAM can effectively highlight crucial temporal features and channels through their convolutional operations, which multihead attention may struggle to achieve. Therefore, this work employs a CBAM-like attention block to better adaptively highlight vital channels and data samples while suppressing less important ones.

Instead of utilizing the entire raw data, this article segregates the input signal into two identical-dimension signals corresponding to distinct frequency bandwidths: high- and low-frequency bands. These high- and low-frequency sEMG signals are fed into the model in parallel. Subsequently, an attention block is applied to the concatenated output. This approach allows us to initially extract frequency-band-specific features, followed by a focus on specific features extracted from these bands. Such a process enables the model to learn detailed features from different frequency regions while prioritizing the most critical attributes. We also employ a similar attention mechanism to the kernel and hidden layer sizes. Our model employs five different kernel sizes and three hidden layers for each kernel. All these outputs are then concatenated and supplied to an attention mechanism, which adaptively underscores the most significant outputs from the hidden layer and kernel.

The frequency-split feature extraction method employs diverse kernel sizes tailored to distinct frequency bands in sEMG signals. This strategy is instrumental in extracting features that not only offer versatility but also improve the overall performance of the model in capturing variations in lower limb motions and hand gestures. The selective application of various kernel sizes to specific frequency bands plays a crucial role in capturing local spectral features, thereby enhancing the model's adaptability and precision. Our approach, distinct from prior methods, underscores our commitment to developing a model that adeptly addresses the intricate dynamics of both lower limb and hand gesture movements. It reflects our focused effort on generalizing the model to be highly effective across different sEMG dataset variations, including sensor types, subject demographics, and muscle groups.

The primary contributions of this article are outlined as follows.

- 1) We apply a motion classification model to a dataset collected from professional football players, marking a unique addition to sEMG-based sports science evaluations.
- 2) Our model outperforms the state-of-the-art by achieving a 3.29% accuracy improvement on the Ninapro DB2 dataset and an 8.66% improvement on the DB2-B dataset.
- 3) For the first time, a model developed for HGR is tested on a soccer dataset that includes lower extremity muscles, professional athletes, and a different sensor type.

The model performed with 96.8% accuracy without any parameter alteration.

- 4) We propose using the 1-D CBAM instead of the conventional MHA for sEMG applications to emphasize vital features in both channel and temporal dimensions.
- 5) We present frequency-split feature extraction, an innovative method in EMG motion classification. This approach processes sEMG data through parallel models across different frequency bands, enabling the model to selectively concentrate on frequency-specific attributes.
- 6) The proposed model is universal for all sEMG-based motion classification tasks. It employs five kernel sizes, three depth sizes, and an attention mechanism. This configuration enhances the model's versatility, allowing it to adapt to various sEMG-based tasks without adjustments to the kernel or hidden layer sizes.

The organization of this article is as follows. A detailed description of the dataset preparation and methodology is presented in Section II. Following this, the specifics of the experiments and data collection are explored in Section III. The results are then evaluated in Section IV. Finally, a discussion of the results and the concluding remarks are provided in Section V.

## II. METHODOLOGY

### A. Database

In this study, we conduct experiments using two datasets: Ninapro, denoted as  $\mathcal{D}_n$ , and Soccer, denoted as  $\mathcal{D}_s$ . Each dataset consists of pairs  $(x^i, y^i)$ , where  $x^i$  represents the recorded sEMG signal,  $y^i$  signifies the true label, and  $i$  is the index of the data sample. The details of dataset  $\mathcal{D}_n$  and  $\mathcal{D}_s$  are discussed in Sections II-A1 and II-A2, respectively.

**1) Ninapro  $\mathcal{D}_n$ :** The Ninapro dataset,  $\mathcal{D}_n$ , is a recognized benchmark for evaluating motion classification models. This study concentrates on the subsets DB2,  $\mathcal{D}_n$ , and DB2-B,  $\mathcal{D}_{n,b}$ , due to their frequent use in contemporary research for model analysis.  $\mathcal{D}_n$  includes data from 49 hand gestures obtained from 40 healthy adults through 12 Delsys Trigno sEMG sensors operating at a 2-kHz sampling rate. Each gesture in  $\mathcal{D}_n$  consists of six 5-s isometric hold repetitions. Following standard practice, we train the model using repetitions 1, 3, 4, and 6 and test using repetitions 2 and 5.  $\mathcal{D}_{n,b}$  is a subset of  $\mathcal{D}_n$ , which includes 17 isotonic and isometric hand and wrist gestures.

**2) Soccer ( $\mathcal{D}_s$ ):** The Soccer dataset,  $\mathcal{D}_s$ , consists of nine, distinct motion types collected from 45 elite players of two EPL football clubs. Exclusively gathered for this study,  $\mathcal{D}_s$  was recorded using Neurocess sEMG sensors, utilizing eight channels at a 1-kHz sampling rate. Specifics regarding the types of movements, sensor placements, and dataset size are detailed in Section III.

For both datasets  $\mathcal{D}_s$  and  $\mathcal{D}_n$ , we aggregate data from all subjects during both the training and testing phases. This methodology ensures that our analysis is *subject-independent*, allowing us to generalize the performance of our model across a diverse range of individuals. By aggregating data from multiple subjects, we aim to capture the variability

present in real-world scenarios, enhancing the robustness and applicability of our proposed approach.

### B. Preprocessing

Several motion classification models have exhibited enhanced performance when sEMG data are preprocessed, as indicated in [22], [29], and [38]. In our approach, we follow to the preprocessing steps outlined in [39], which include Butterworth filtering, normalization, and segmentation. The goal of this article is to improve classification performance through the application of an attention mechanism to both high- and low-frequency bands. To this end, we introduce two stages of preprocessing to obtain preprocessed sEMG with high-frequency bands, denoted as  $x_h$ , and a low-frequency band,  $x_l$ . The proposed frequency band split strategy involves equalizing the power spectral density among the sampled signals from different datasets. The low-frequency band is defined as 0–100 Hz, whereas the high-frequency band ranges from 100 to 300 Hz. This division is based on the finding that the 300–500-Hz band holds a minimal proportion (about 0.02) of the total power in sEMG samples. In addition, the median frequency for power spectral density in our dataset approximates 96 Hz. This aligns with literature suggesting a median sEMG frequency of 80–120 Hz, variable with subject demographics [40], [41]. The emphasis on equalizing power spectral density ensures that our twin network architecture processes each band effectively, maintaining a balance in signal complexity for efficient feature extraction. By segmenting the signal into these specific frequency bands, our system enhances its analytical capability and feature extraction efficiency, especially in isolating distinct aspects of the signal.

**1) Preprocessing sEMG With High-Frequency Band  $x_h$ :** Initially, we apply a ninth-order bandpass Butterworth filter with cutoff frequencies at 100 and 300 Hz. Following this, we implement a  $\mu$ -law transformation, as defined in (1), on the filtered signal, setting  $\mu = 2048$ :

$$x_\mu = \text{sign}(x) \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)}. \quad (1)$$

After the  $\mu$ -law transformation, we apply minmax normalization as given in the following:

$$x_{\text{norm}} = \frac{x_\mu - x_{\mu,\min}}{x_{\mu,\max} - x_{\mu,\min}}. \quad (2)$$

After normalization, the signal is segmented using the sliding window technique with a window size and overlap provided in Table I.

**2) Preprocessing sEMG With Low-Frequency Band  $x_l$ :** Initially, we utilize a ninth-order Butterworth low-pass filter to eliminate high-frequency components exceeding 100 Hz. For further data smoothing, we implement a Savitzky–Golay filter with a window size of  $W_n = 51$  and polynomial order of  $P_o = 3$ , as demonstrated in [42]. The Savitzky–Golay filter is adept at preserving notable features of the underlying signal while providing a smoother data estimate [43]. Following the filtering process, the same normalization and segmentation steps are applied as detailed in Section II-B1.

**TABLE I**  
DETAILS OF PREPROCESS STEPS FOR EACH DATASET  
USED IN THIS ARTICLE

	$\mathcal{D}_n$	$\mathcal{D}_{n,b}$	$\mathcal{D}_s$
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

In Table I, the preprocessing steps are detailed, along with the sizes of the training and test sets for each dataset, namely,  $\mathcal{D}_n$ ,  $\mathcal{D}_{n,b}$ , and  $\mathcal{D}_s$ .

### C. Proposed Model

The proposed model, henceforth referred to as convolutional octave-band zooming-in with depth-kernel attention learning (COZDAL), employs an attention mechanism that targets various frequency bands, kernel sizes, and hidden layer depths. The acronym “COZDAL” aptly captures these key characteristics.

**1) Network Architecture:** The architecture of the proposed model, shown in Fig. 1, accepts  $x_h$  and  $x_l$  as 3-D inputs, with a shape corresponding to batch size, channel number, and window length. Each signal independently proceeds through a KD block for feature extraction. The outputs from the KD block are unified along the channel dimensions via the concatenation block, resulting in a merged output. This is then passed to a 1-D convolution block with kernel size and stride of 1, followed by an attention block emphasizing the model’s focus on key channels and samples.

Following the attention block, the model navigates through various adaptive average pooling layers and convolution blocks. This process is purposed to decrease the signal’s channel size and sequence length to align with the number of prediction classes.

The application of distinct KD blocks for the signal’s high- and low-frequency bands is motivated by the intent to harness the model’s attention toward differing frequency bandwidths. After the channelwise concatenation and 1-D convolution, the attention block learns to focus on the features extracted from specific frequency bands.

**2) KD Block:** The KD block is designed to improve the performance of deep neural networks for time series feature learning. This is achieved by using 1-D convolution blocks with a variety of kernel sizes and depths. However, finding the right model depth and kernel size can be a challenging task, often requiring numerous trials to identify the optimal parameters. Furthermore, these optimal parameters differ depending on the dataset and application.

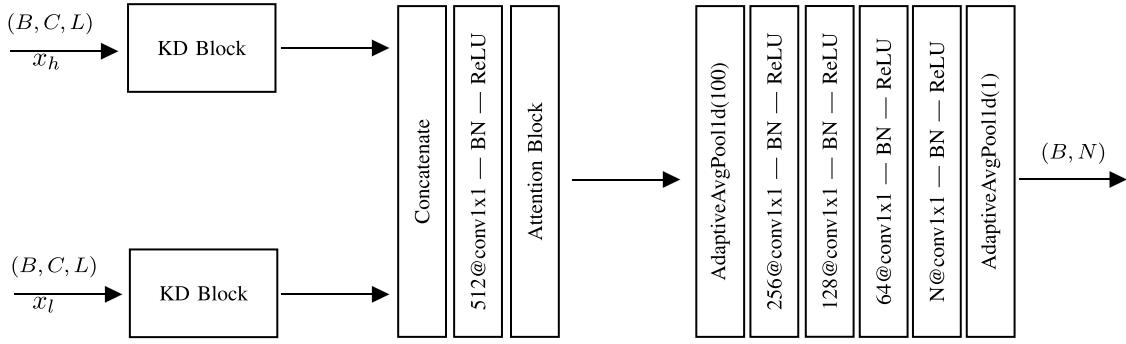
Recent studies have indicated that using multiple kernel sizes simultaneously can enhance model performance [44],

[45]. This is due to each kernel size’s ability to capture unique data features. Therefore, our design incorporates five different kernels through the D block architecture, detailed in Section II-C3. We combine all D block outputs and a residual tensor along the channel dimension, which allows the model to capture unique features from different kernels. The residual tensor is generated from a sequence of a max-pooling layer and a 1-D convolution operation with a kernel size of 1.

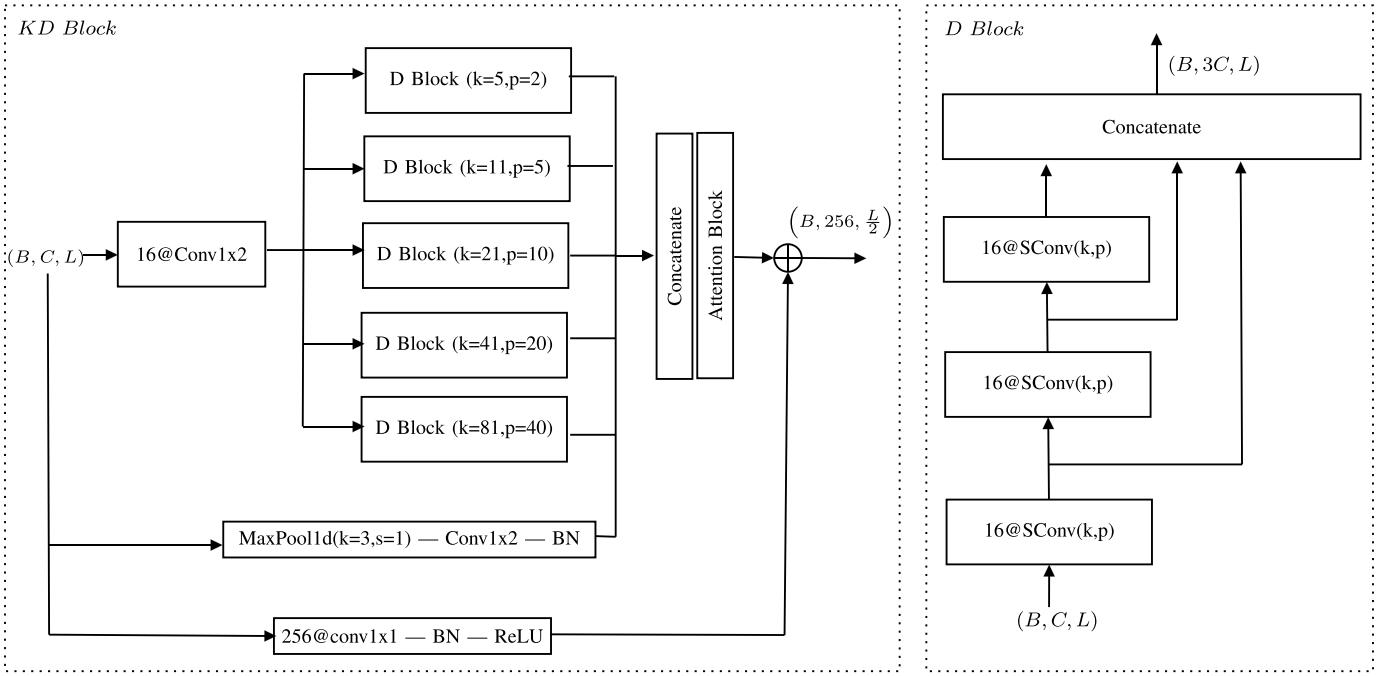
As shown in Fig. 2, the model first goes through a 1-D convolution block with a kernel size of 1 before utilizing D blocks of various kernels. This step adjusts the signal length, i.e., downsampling via stride, and channel size, i.e., output filter size. After combining the D block outputs along the channel dimension, the model is fed into an attention block. This allows the model to focus on the most crucial kernel size and depth. The output from the attention block is combined with a residual via elementwise summation. The integration of the residual serves two critical functions—mitigating the vanishing gradient problem, which can hinder the learning process in deep networks, and augmenting feature reusability across layers [46]. The residual is obtained through a 1-D convolution block with a stride of 2, batch normalization, and output filter size equal to the total number of channels after concatenation. The KD block accepts an input of shape  $(B, C,$  and  $L)$ , representing batch size, channel count, and window length, respectively, and outputs a tensor of shape  $(B, 16C,$  and  $\frac{L}{2})$ .

**3) D Block:** The D block is engineered to assess the performance enhancement resulting from varying depths, specifically cascaded layers of 1-D separable convolution blocks (SConvs). The adoption of SConv, as opposed to traditional convolutional layers, promotes computational and parameter efficiency, mitigates overfitting potential, and elevates feature extraction due to the independent handling of spatial and channelwise features [47]. As shown in Fig. 2, the D block applies a sequence of SConvs with uniform kernel size and padding. Each SConv block’s output is concatenated along the channel dimension. Instead of solely using the final SConv block’s result, we incorporate the output of each depth layer into the final output, aiming to utilize the residual connection effect. This approach helps to counteract vanishing gradients and integrates valuable features from each depth level. Each level, i.e., SConv output, carries unique features that may collectively enhance model performance. When the output of each SConv is concatenated and subsequent attention in the KD block is applied, the model learns to prioritize depth values that significantly influence classification performance. The input and output filter sizes of SConvs in D block are fixed to 16. As a result of the concatenation, the output channel size of the D block is tripled.

**4) Attention Block:** The attention block in our model incorporates two successive attention modules: channel and temporal attention. Inspired by the CBAM proposed in [37], our design applies 1-D channel and temporal attention to enhance the classification performance for sequential data. The channel attention module highlights the most influential channels, while the temporal attention module emphasizes the significant samples across the sequence length. Given the input



**Fig. 1.** Schematic of the proposed model is provided. The model accepts two input tensors of shape  $\triangleright B$ ,  $C$ , and  $L$ , corresponding to low-frequency band,  $x_h$ , and high-frequency band,  $x_l$ .  $B$  corresponds to the batch size,  $C$  refers to the number of channels, and  $L$  indicates the sequence length. Convolution layers are denoted as  $f@Convkxs$ , 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  $\mathcal{D}_n$ , 17 for  $\mathcal{D}_{n,b}$ , and 9 for  $\mathcal{D}_s$ .



**Fig. 2.** Schematics of KD block (left) and D block (right) are provided. KD block consists of an initial 1-D 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  $(B, C$ , and  $L)$  into an output tensor of shape  $(B, 256$ , and  $\frac{L}{2}$ ). D block illustrates the cascaded 1-D SConvs operation, starting with an input dimension of  $(B, C$ , and  $L$ ). The outputs of all SConv blocks are concatenated, tripling the channel dimension, resulting in the final shape  $(B, 3C$ , and  $L$ ). Both structures collaboratively optimize feature learning, exploiting various depths and kernel sizes to enhance the model's efficiency and performance.

feature tensor  $h$  to the attention block, the output  $h'$  can be calculated as

$$h' = F_t(F_{ch}(h) \otimes h) \otimes (F_{ch}(h) \otimes h) \quad (3)$$

where  $F_{ch}(\cdot)$  and  $F_t(\cdot)$  denote the channel and temporal attention functions, respectively, and  $\otimes$  represents the matrix multiplication. The channel attention function,  $F_{ch}(\cdot)$ , is defined as

$$F_{ch}(h) = \sigma(F_{mlp}(h_{avg}) + F_{mlp}(h_{max})) \quad (4)$$

where  $h_{avg}$  and  $h_{max}$  are tensors obtained by performing 1-D average pooling and 1-D max pooling over the temporal dimension of  $h$ , respectively. Here,  $\sigma$  denotes the sigmoid activation function.  $F_{mlp}(\cdot)$  signifies a multilayer perceptron with a

single hidden layer. Drawing from the squeeze-and-excitation network approach in [48], we employ  $F_{mlp}(\cdot)$  to enhance attention. Initially, the number of perceptrons in  $F_{mlp}(\cdot)$  is reduced by a specified reduction ratio before reverting to the input shape corresponding to the channel number. The reduction ratio is selected as 16 in the KD block and 128 in the final attention module, as shown in Fig. 1.

The temporal attention function,  $F_t(\cdot)$ , operates by initially computing the mean and max values of the input tensor across the channel dimension. As a result, two tensors are produced, each representing the maximum and mean values across the sequence length of all channels. These tensors are then concatenated along the channel dimension, creating a new tensor that is subsequently fed into a 1-D convolution block.

**Algorithm 1** 1-D Attention Block Module

```

1: procedure CHANNEL ATTENTION( $x$ )
2:    $x_{\text{avg}} \leftarrow \text{AdaptiveAvgPool1d}(x)$ 
3:    $x_{\text{max}} \leftarrow \text{AdaptiveMaxPool1d}(x)$ 
4:    $x_{\text{avg}} \leftarrow \text{mlp}(x_{\text{avg}})$ 
5:    $x_{\text{max}} \leftarrow \text{mlp}(x_{\text{max}})$ 
6:    $x_{\text{sum}} \leftarrow x_{\text{avg}} + x_{\text{max}}$ 
7:    $x_{\text{sig}} \leftarrow \text{sigmoid}(x_{\text{sum}}) \otimes x$ 
8:   return  $x_{\text{sig}}$ 
9: end procedure
10:
11: procedure TEMPORAL ATTENTION( $x$ )
12:    $x_{\text{avg}} \leftarrow \text{mean}(x, \text{dim} = 1, \text{keepdim} = \text{True})$ 
13:    $x_{\text{max}} \leftarrow \text{max}(x, \text{dim} = 1, \text{keepdim} = \text{True})$ 
14:    $x_{\text{cat}} \leftarrow \text{concatenate}(x_{\text{avg}}, x_{\text{max}}, \text{dim} = 1)$ 
15:    $x_{\text{conv}} \leftarrow \text{conv}(x_{\text{cat}})$ 
16:    $x_{\text{sig}} \leftarrow \text{sigmoid}(x_{\text{conv}}) \otimes x$ 
17:   return  $x_{\text{sig}}$ 
18: end procedure
19:
20: procedure ATTENTION BLOCK( $x$ )
21:    $x \leftarrow \text{Channel Attention}(x)$ 
22:    $x \leftarrow \text{Temporal Attention}(x)$ 
23:   return  $x$ 
24: end procedure

```

This block has an output filter size of 1, a kernel size of 3, and a padding of 1, effectively focusing on salient temporal features while maintaining the spatial context. Algorithm 1 offers a detailed pseudocode depiction of the attention module's implementation.

### III. EXPERIMENTS

The sEMG data were recorded from 45 male first-team athletes of two EPL clubs via the Neurocess system, which consisted of wireless, dry, active, and bipolar sEMG sensors [49]. During the experiments, a total of eight sEMG sensors were positioned on each subject's biceps femoris long head, semitendinosus, adductor longus, and soleus muscles bilaterally in line with the SENIAM standards [50]. For the repeatability of the experiments, we detailed the sensor positioning on each muscle group as follows.

- 1) *Biceps Femoris Long Head*: The sensors were placed at 50% on the line between the ischial tuberosity and the lateral epicondyle of the tibia.
- 2) *Semitendinosus*: The sensors were placed at 50% on the line between the ischial tuberosity and the medial epicondyle of the tibia.
- 3) *Adductor Longus*: The sensors were placed at 66% on the line from the pubic tubercle to the medial epicondyle of the femur, over the muscle belly.
- 4) *Soleus*: The sensors were placed at 66% of the line between the medial condylis of the femur to the medial malleolus.

The athletes were instructed to perform the following hamstring exercises.

- 1) *60° Adductor Squeeze*: The athlete lay on their back with 60° hip flexion and squeezed a foam roll between the knees.
- 2) *Hamstring Claw*: The athlete lay face down over a box, hips at 45° flexion, knees at 15° flexion, and extended the thigh while maintaining posture.
- 3) *Hip Extension*: The athlete lay face down, pelvis on a mat, hips at 30° flexion, knees at 60° flexion, and extended the hip to lift the thigh.
- 4) *Hip Flexion*: The athlete lay on their back, hips at 60°, feet flat, and lifted the test leg's heel from the floor.
- 5) *Sitting Soleus Raise*: The athlete sat on a 45-cm box, knee at 90°, and tried to raise the heel with maximal force.
- 6) *Single Leg (SL) Bent-Knee Soleus Raise*: The athlete stood on one leg, knee bent at 30°, and raised the heel.
- 7) *SL Elevated Glute Bridge*: The athlete lay on their back, knees at 90°, feet on a 45 cm box, and lifted the hip off the floor.
- 8) *Isometric Prone Squeeze at 0°*: The athlete lay face down with the knee straight. An examiner provided resistance at the ankle, and the athlete tried to contract against it.
- 9) *SL Prone Curls*: The athlete lay face down, knee straight, and then curled the knee to 90° and back, keeping the hip stable.

In this research, experiments were performed on three datasets:  $\mathcal{D}_n$ ,  $\mathcal{D}_{n,b}$ , and  $\mathcal{D}_s$ , using accuracy as the evaluation metric. The experimental protocols and sensor placements were supervised by the clubs' lead physiotherapists, with sEMG recordings conducted at the training grounds under the ethical clearance of King's College London (reference MRSP-20/21-20996). All personal information was anonymized, adhering to privacy guidelines. Each dataset was split into training, validation, and test sets, with details provided in Table I. The training set was used to learn the model, the validation set to prevent overfitting, and the test set for comparing methods. We ran the experiments ten times with different random seeds and averaged the results.

### IV. RESULTS

The COZDAL model is benchmarked against state-of-the-art models on the Ninapro datasets  $\mathcal{D}_n$  and  $\mathcal{D}_{n,b}$ . Accuracy scores of each model are summarized in Table II. To ensure a fair comparison, identical training and testing repetitions and a consistent window size of 200 ms are used as outlined in Section II-A1. No parameter tuning is conducted on our proposed 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  $\mathcal{D}_n$ , 17 for  $\mathcal{D}_{n,b}$ , and 9 for  $\mathcal{D}_s$ .

For the dataset  $\mathcal{D}_n$ , COZDAL achieves a classification accuracy of 95.30%, surpassing the second-best contemporary model by 3.29%. When comparing COZDAL's performance to the collection of seven listed contemporary works, its accuracy score exceeds the mean by 15.7% and the median by 13.1%, reflecting a significant advancement in the state-of-the-art.

TABLE II

COMPARISON OF THE COZDAL MODEL'S ACCURACY WITH OTHER STATE-OF-THE-ART MODELS ACROSS THE  $\mathcal{D}_n$ ,  $\mathcal{D}_{n,b}$ , AND  $\mathcal{D}_s$  DATASETS, EVALUATED AT A 200-ms WINDOW SIZE

Model	Dataset	Accuracy (%) $\pm$ Std.	F1 Score
CNN by Aztori et al. [22]	$\mathcal{D}_n$	60.30	-
Feature Engineering [51]	$\mathcal{D}_n$	75.30	-
CNNLM [52]	$\mathcal{D}_n$	78.71	-
CNN-RNN by Hu et al. [32]	$\mathcal{D}_n$	82.20	-
CNN by Wei et al. [53]	$\mathcal{D}_n$	83.70	-
TraHGR-Huge [35]	$\mathcal{D}_n$	86.18 $\pm$ 4.99	-
EMGHandNet [33]	$\mathcal{D}_n$	92.01	-
<b>COZDAL</b>	$\mathcal{D}_n$	<b>95.30 <math>\pm</math> 2.77</b>	<b>0.95</b>
SVM by Zhai et al. [54]	$\mathcal{D}_{n,b}$	81.07	-
CNN by Zhai et al. [54]	$\mathcal{D}_{n,b}$	82.22	-
CNN by Ding et al. [55]	$\mathcal{D}_{n,b}$	83.79	-
TraHGR-Huge [35]	$\mathcal{D}_{n,b}$	88.91	-
CNN by Rahimian et al. [34]	$\mathcal{D}_{n,b}$	90.14	0.92
<b>COZDAL</b>	$\mathcal{D}_{n,b}$	<b>98.80 <math>\pm</math> 2.24</b>	<b>0.99</b>
<b>COZDAL</b>	$\mathcal{D}_s$	<b>96.30 <math>\pm</math> 3.18</b>	<b>0.97</b>

TABLE III

SUMMARY OF CLASSIFICATION METRICS FOR A MULTICLASS PREDICTION MODEL, DISPLAYING PRECISION, RECALL, AND F1-SCORE FOR NINE CLASSES IN  $\mathcal{D}_s$

	precision	recall	f1-score	support
1	0.95	0.97	0.96	2290
2	1.00	1.00	1.00	2358
3	0.99	1.00	1.00	2323
4	1.00	1.00	1.00	2288
5	1.00	1.00	1.00	2241
6	0.96	0.95	0.95	2822
7	0.94	0.92	0.93	3376
8	0.97	0.98	0.97	2417
9	0.93	0.93	0.93	3109
accuracy			0.97	23224
macro avg	0.97	0.97	0.97	23224
weighted avg	0.97	0.97	0.97	23224

Within the dataset  $\mathcal{D}_{n,b}$ , COZDAL manifests a remarkable accuracy of 98.80%, distinctly setting a new benchmark. This score not only exceeds the closest competitor, CNN by Rahimian et al. [34], by 8.66% but also surpasses the mean and median of the other contemporary models listed by 13.57% and 15.01%, respectively.

COZDAL attains an accuracy of 96.30% in classifying nine hamstring exercises within the dataset  $\mathcal{D}_s$ . The corresponding classification metrics, including precision, recall, and F1-score, are presented for the nine classes in Table III, with exercise numbers consistent with those described in Section III. Notably, the model achieves a 100% F1 score for exercises 2–5, reflecting their isometric nature with minor isotonic contractions at the onset and conclusion. In contrast, exercises 6, 7, and 9, characterized by full isotonic movements without isometric holds, yield relatively lower F1 scores. This difference is attributed to the more dynamic nature of isotonic

TABLE IV

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
accuracy			0.99	60259
macro avg	0.99	0.99	0.99	60259
weighted avg	0.99	0.99	0.99	60259

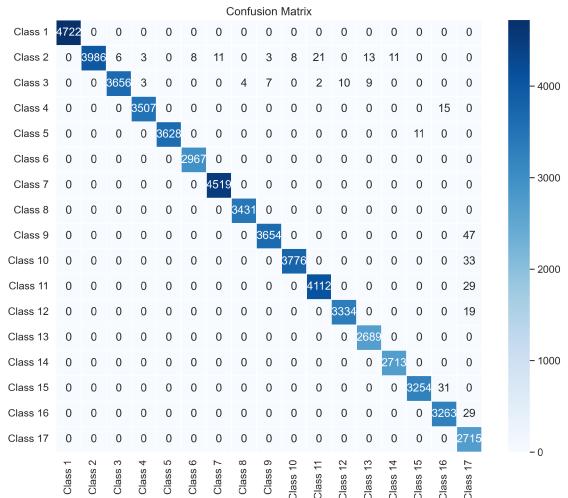
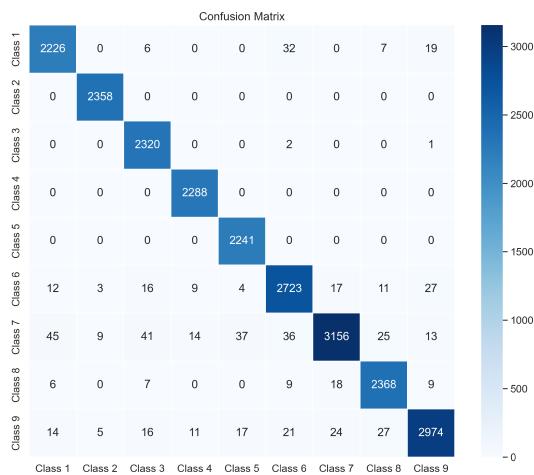


Fig. 3. Confusion matrix showcasing the classification performance across 17 classes of  $\mathcal{D}_{n,b}$ . The diagonal elements represent the number of correct predictions for each class, while off-diagonal elements are misclassifications. The color gradient reflects the magnitude of instances, illustrating the model's precision in classification tasks.

contractions, where the chosen 200-ms window is insufficient for discriminating all instances.

The classification performance of our model COZDAL is visually represented through confusion matrices for both  $\mathcal{D}_{n,b}$  and  $\mathcal{D}_s$ , as shown in Figs. 3 and 4, respectively. The confusion matrices provide a comprehensive view of the model's predictive accuracy across multiple classes.

Fig. 3 and Table IV show the performance over the 17 classes of  $\mathcal{D}_{n,b}$ . In Fig. 3, the diagonal elements, which



**Fig. 4.** Confusion matrix showcasing the classification performance across nine classes of  $\mathcal{D}_s$ . The diagonal elements represent the number of correct predictions for each class, while off-diagonal elements are misclassifications. The color gradient reflects the magnitude of instances, illustrating the model's precision in classification tasks.

represent correct predictions for each class, indicate high accuracy. The off-diagonal elements, corresponding to misclassifications, are significantly lower, showcasing the model's effectiveness in accurately classifying instances across different classes. The color gradient in the matrix further aids in visualizing the magnitude of correct and incorrect classifications, thereby indicating uniform precision in classification tasks.

Similarly, Fig. 4 presents the confusion matrix for  $\mathcal{D}_s$ , encompassing nine classes. This matrix also demonstrates a high rate of correct predictions, as evident from the prominent diagonal elements. This uniformity in the classification accuracy across different classes suggests that the model does not exhibit a bias toward any specific class, maintaining a uniformly high predictability rate across all classes.

## V. CONCLUSION AND DISCUSSION

In this work, we present the COZDAL model, which demonstrates substantial advancement over contemporary models across multiple datasets, including  $\mathcal{D}_n$  and  $\mathcal{D}_{n,b}$ . The model achieves superior accuracies of 95.30% and 98.80% for  $\mathcal{D}_n$  and  $\mathcal{D}_{n,b}$ , respectively, surpassing leading contemporary models by 3.29% and 8.66%.

A notable advantage of the COZDAL model is its innovative integration of channel and temporal attention mechanisms. These mechanisms enable the model to selectively focus on the most relevant feature vectors derived from convolution blocks of various kernel sizes and frequency bands. This focus not only enhances model performance by emphasizing relevant time steps in the data but also improves adaptation to dataset variations.

Moreover, the employment of various kernel sizes in the COZDAL model allows for the extraction of a diverse range of signal features. Each kernel size is adept at isolating unique signal characteristics, enhancing the model's performance and versatility. This approach is particularly effective in capturing high- and low-frequency components, further augmenting the model's accuracy and adaptability.

In addition, our frequency-split feature extraction technique significantly contributes to the model's versatility and precision. By concentrating on specific frequency bands during feature extraction, the COZDAL model learns to better identify features unique to each band, thereby adapting more effectively to data variations. Similar to the kernel size selection, splitting the data into distinct high- and low-frequency bands improves the model's feature extraction capabilities.

The COZDAL model's ability to adapt to different datasets without extensive parameter tuning is another key advantage, enhancing its applicability in diverse scenarios. This adaptability is evidenced through extensive testing, where only the final output layer size is modified to suit the number of classes in each dataset.

Applying the COZDAL model to datasets from elite-level soccer athletes represents a significant milestone in sports science and injury prevention. This application is the first instance of using sEMG-based deep learning models on data from professional athletes for motion classification, marking a significant step toward realizing the full potential of data-driven sports applications.

Traditional motion classification models, often limited by their design for specific muscle groups, have shown decreased accuracy in real-life sports applications due to inconsistent muscle group selection and varying activation patterns. The COZDAL model addresses these challenges, demonstrating robust performance across different muscle groups and subject demographics.

Looking ahead, we see promising opportunities in combining feature learning with meta-learning or transfer learning to further address the challenges in sEMG-based motion classification within sports applications. Future research will explore these methods, potentially unlocking significant advancements in the field and bridging the gap between current limitations and the practical needs of sports-based motion classification.

## REFERENCES

- [1] M. Ergeneci, K. Gokcesu, E. Ertan, and P. Kosmas, "An embedded, eight channel, noise canceling, wireless, wearable sEMG data acquisition system with adaptive muscle contraction detection," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 1, pp. 68–79, Feb. 2018.
- [2] T. Kamali, R. Boostani, and H. Parsaei, "A multi-classifier approach to MUAP classification for diagnosis of neuromuscular disorders," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 191–200, Jan. 2014.
- [3] S. S. Nair, R. M. French, D. Laroche, and E. Thomas, "The application of machine learning algorithms to the analysis of electromyographic patterns from arthritic patients," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 174–184, Apr. 2010.
- [4] B. Larsson, S. Karlsson, M. Eriksson, and B. Gerdle, "Test-retest reliability of EMG and peak torque during repetitive maximum concentric knee extensions," *J. Electromogr. Kinesiol.*, vol. 13, no. 3, pp. 281–287, Jun. 2003.
- [5] K. Gokcesu, M. Ergeneci, E. Ertan, A. Z. Alkilani, and P. Kosmas, "An sEMG-based method to adaptively reject the effect of contraction on spectral analysis for fatigue tracking," in *Proc. ACM Int. Symp. Wearable Comput.*, Oct. 2018, pp. 80–87.
- [6] A. Jaramillo-Yáñez, M. E. Benalcázar, and E. Mena-Maldonado, "Real-time hand gesture recognition using surface electromyography and machine learning: A systematic literature review," *Sensors*, vol. 20, no. 9, p. 2467, Apr. 2020.
- [7] J. L. Betthauser et al., "Stable responsive EMG sequence prediction and adaptive reinforcement with temporal convolutional networks," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 6, pp. 1707–1717, Jun. 2020.

- [8] M. Ergeneci, D. Carter, and P. Kosmas, "sEMG onset detection via bidirectional recurrent neural networks with applications to sports science," *IEEE Sensors J.*, vol. 22, no. 19, pp. 18751–18761, Oct. 2022.
- [9] K. Bennell et al., "Isokinetic strength testing does not predict hamstring injury in Australian rules footballers," *Brit. J. Sports Med.*, vol. 32, no. 4, pp. 309–314, Dec. 1998.
- [10] M. Buckthorpe et al., "Recommendations for hamstring injury prevention in elite football: Translating research into practice," *Brit. J. Sports Med.*, vol. 53, no. 7, pp. 449–456, Nov. 2018.
- [11] B. Green, M. N. Bourne, N. van Dyk, and T. Pizzari, "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," *Brit. J. Sports Med.*, vol. 54, no. 18, pp. 1081–1088, Apr. 2020.
- [12] B. C. Heiderscheit, M. A. Sherry, A. Silder, E. S. Chumanov, and D. G. Thelen, "Hamstring strain injuries: Recommendations for diagnosis, rehabilitation, and injury prevention," *J. Orthopaedic Sports Phys. Therapy*, vol. 40, no. 2, pp. 67–81, Feb. 2010.
- [13] R. A. Clarke, "Hamstring injuries: Risk assessment and injury prevention," *Ann. Acad. Med. Singap.*, vol. 37, no. 4, pp. 341–346, Apr. 2008.
- [14] N. Malliaropoulos et al., "Hamstring exercises for track and field athletes: Injury and exercise biomechanics, and possible implications for exercise selection and primary prevention," *Brit. J. Sports Med.*, vol. 46, no. 12, pp. 846–851, Sep. 2012.
- [15] J. Orchard and H. Seward, "Epidemiology of injuries in the Australian football league, seasons 1997–2000," *Brit. J. Sports Med.*, vol. 36, no. 1, pp. 39–44, Feb. 2002.
- [16] M. C. W. Elliott, B. Zarins, J. W. Powell, and C. D. Kenyon, "Hamstring muscle strains in professional football players: A 10-year review," *Amer. J. Sports Med.*, vol. 39, no. 4, pp. 843–850, Apr. 2011.
- [17] J. Ekstrand, J. C. Healy, M. Waldén, J. C. Lee, B. English, and M. Hägglund, "Hamstring muscle injuries in professional football: The correlation of MRI findings with return to play," *Brit. J. Sports Med.*, vol. 46, no. 2, pp. 112–117, Feb. 2012.
- [18] C. Eirale, A. Farooq, F. A. Smiley, J. L. Tol, and H. Chalabi, "Epidemiology of football injuries in asia: A prospective study in Qatar," *J. Sci. Med. Sport*, vol. 16, no. 2, pp. 113–117, Mar. 2013.
- [19] D. Xiong, D. Zhang, X. Zhao, and Y. Zhao, "Deep learning for EMG-based human-machine interaction: A review," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 3, pp. 512–533, Mar. 2021.
- [20] M. Simão, N. Mendes, O. Gibaru, and P. Neto, "A review on electromyography decoding and pattern recognition for human-machine interaction," *IEEE Access*, vol. 7, pp. 39564–39582, 2019.
- [21] K.-H. Park and S.-W. Lee, "Movement intention decoding based on deep learning for multiuser myoelectric interfaces," in *Proc. 4th Int. Winter Conf. Brain-Comput. Interface (BCI)*, Feb. 2016, pp. 1–2.
- [22] M. Atzori, M. Cognolato, and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers NeuroRobotics*, vol. 10, p. 9, Sep. 2016.
- [23] M. Zia ur Rehman et al., "Stacked sparse autoencoders for EMG-based classification of hand motions: A comparative multi day analyses between surface and intramuscular EMG," *Appl. Sci.*, vol. 8, no. 7, p. 1126, Jul. 2018.
- [24] M. Zia ur Rehman et al., "Multiday EMG-based classification of hand motions with deep learning techniques," *Sensors*, vol. 18, no. 8, p. 2497, Aug. 2018.
- [25] S. Karlsson and B. Gerdle, "Mean frequency and signal amplitude of the surface EMG of the quadriceps muscles increase with increasing torque—A study using the continuous wavelet transform," *J. Electromogr. Kinesiol.*, vol. 11, no. 2, pp. 131–140, Apr. 2001.
- [26] D. Farina et al., "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 797–809, Jul. 2014.
- [27] E. Rahimian, S. Zabih, A. Asif, D. Farina, S. F. Atashzar, and A. Mohammadi, "FS-HGR: few-shot learning for hand gesture recognition via electromyography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1004–1015, 2021.
- [28] U. C. Allard et al., "A convolutional neural network for robotic arm guidance using sEMG based frequency-features," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 2464–2470.
- [29] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Sci. Rep.*, vol. 6, no. 1, p. 36571, Nov. 2016.
- [30] Y. Du, W. Jin, W. Wei, Y. Hu, and W. Geng, "Surface EMG-based inter-session gesture recognition enhanced by deep domain adaptation," *Sensors*, vol. 17, no. 3, p. 458, Feb. 2017.
- [31] U. Cote-Allard et al., "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Apr. 2019.
- [32] Y. Hu, Y. Wong, W. Wei, Y. Du, M. Kankanhalli, and W. Geng, "A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition," *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0206049.
- [33] N. K. Karnam, S. R. Dubey, A. C. Turlapati, and B. Gokaraju, "EMGHandNet: A hybrid CNN and bi-LSTM architecture for hand activity classification using surface EMG signals," *Biocybernetics Biomed. Eng.*, vol. 42, no. 1, pp. 325–340, Jan. 2022.
- [34] E. Rahimian, S. Zabih, S. F. Atashzar, A. Asif, and A. Mohammadi, "Semb-based hand gesture recognition via dilated convolutional neural networks," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2019, pp. 1–5.
- [35] S. Zabih, E. Rahimian, A. Asif, and A. Mohammadi, "TraHGR: Transformer for hand gesture recognition via ElectroMyography," 2022, *arXiv:2203.16336*.
- [36] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.
- [37] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.
- [38] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, and W. Geng, "A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface," *Pattern Recognit. Lett.*, vol. 119, pp. 131–138, Mar. 2019.
- [39] E. Rahimian, S. Zabih, A. Asif, D. Farina, S. F. Atashzar, and A. Mohammadi, "FS-HGR: Few-shot learning for hand gesture recognition via electromyography," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1004–1015, 2021.
- [40] W. H. J. P. Linssen, D. F. Stegeman, E. M. G. Joosten, S. L. H. Notermans, M. A. van't Hof, and R. A. Binkhorst, "Variability and interrelationships of surface EMG parameters during local muscle fatigue," *Muscle Nerve*, vol. 16, no. 8, pp. 849–856, Aug. 1993.
- [41] M. M. Lowery, C. L. Vaughan, P. J. Nolan, and M. J. O'Malley, "Spectral compression of the electromyographic signal due to decreasing muscle fiber conduction velocity," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 3, pp. 353–361, Sep. 2000.
- [42] M. Ergeneci, E. Bayram, A. Y. Yildiz, D. Carter, and P. Kosmas, "sEMG motion classification via few-shot learning with applications to sports science," *Authorea Preprints*, 2023.
- [43] R. W. Schafer, "What is a Savitzky-Golay filter? [Lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 111–117, Jul. 2011.
- [44] F. Chen, J. Wei, B. Xue, and M. Zhang, "Feature fusion and kernel selective in inception-v4 network," *Appl. Soft Comput.*, vol. 119, Apr. 2022, Art. no. 108582.
- [45] H. I. Fawaz et al., "InceptionTime: Finding AlexNet for time series classification," *Data Mining Knowl. Discovery*, vol. 34, no. 6, pp. 1936–1962, Nov. 2020.
- [46] S. Zagoruyko and N. Komodakis, "Wide residual networks," 2016, *arXiv:1605.07146*.
- [47] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1800–1807.
- [48] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
- [49] Neurocess. Accessed: Jun. 2, 2023. [Online]. Available: <https://www.neurocess.co/>
- [50] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, "Development of recommendations for SEMG sensors and sensor placement procedures," *J. Electromogr. Kinesiol.*, vol. 10, no. 5, pp. 361–374, Oct. 2000.
- [51] M. Atzori et al., "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Sci. Data*, vol. 1, no. 1, pp. 1–13, Dec. 2014.
- [52] X. Chen, Y. Li, R. Hu, X. Zhang, and X. Chen, "Hand gesture recognition based on surface electromyography using convolutional neural network with transfer learning method," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 4, pp. 1292–1304, Apr. 2021.

- [53] W. Wei, Q. Dai, Y. Wong, Y. Hu, M. Kankanhalli, and W. Geng, "Surface-electromyography-based gesture recognition by multi-view deep learning," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 10, pp. 2964–2973, Oct. 2019.
- [54] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprostheses control based on convolutional neural network," *Frontiers Neurosci.*, vol. 11, p. 379, Jul. 2017.
- [55] Z. Ding, C. Yang, Z. Tian, C. Yi, Y. Fu, and F. Jiang, "SEMG-based gesture recognition with convolution neural networks," *Sustainability*, vol. 10, no. 6, p. 1865, Jun. 2018.

**Mert Ergeneci** received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2016. He is currently pursuing the Ph.D. degree in telecommunications with King's College London, London, U.K.

He is the Co-Founder of Neurocess Ltd., London, focusing on injury prevention via surface electromyography data analysis in professional sports. His research interests include surface electromyography signal processing and deep learning.

**Erkan Bayram** received the B.S. degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2021. He is currently pursuing the Ph.D. degree in decision and control systems with the Electrical and Computer Engineering Department, University of Illinois at Urbana–Champaign, Champaign, IL, USA.

He is a Consultant at Neurocess Ltd., London, focusing on injury prevention via surface electromyography data analysis in professional sports. His current research interests include decision-making, sensor networks, and signal processing for control systems.

**David Binningsley** received the B.Sc. degree (Hons.) in physiotherapy in 1998.

He joined the famed Middlesbrough FC Academy, Middlesbrough, U.K., before spending a significant tenure of 16 years at Sunderland AFC first team, Sunderland, U.K., and Manchester United, Manchester, U.K., in 2018, where he is the First Team Physiotherapist, focusing primarily on rehabilitating first-team players. In collaboration with Neurocess Ltd., London, U.K., he is exploring the application of surface electromyography to prevent injuries and monitor training loads in football.

**Daryl Carter** received the B.Sc. degree in sports rehabilitation and the M.Sc. degree in physiotherapy from York St John University, York, U.K., in 2009 and 2015, respectively.

He is currently working in professional football with the Leeds United Football Club (FC), Leeds, U.K., as a First Team Physiotherapist. Alongside Neurocess Ltd., London, U.K., he is investigating the use of surface electromyography for injury prevention and the monitoring of training load in football.

**Panagiotis Kosmas** (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Northeastern University, Boston, MA, USA, in 2005.

He joined Kings College London (KCL), London, U.K., as a Lecturer, in 2008, where he is currently a Reader with the Department of Engineering. Prior to his appointment at KCL, he held research positions at the Center for Subsurface Sensing and Imaging Systems (CenSSIS), Boston; the University of Loughborough, Loughborough, U.K.; and the Computational Electromagnetics Group, University of Wisconsin-Madison, Wisconsin, MI, USA. He is also a Co-Founder of Mediwise Ltd., London, an award-winning small and medium-sized enterprise (SME) focusing on the use of electromagnetic waves for medical applications, which is now part of Meta Materials Inc., Dartmouth, NS, Canada. His research interests include microwave engineering with application to sensing and imaging, antenna and sensor design, physics-based detection methods, and inverse problems theory and techniques. He has contributed over 150 journals and conference publications in these areas.

Dr. Kosmas is serving as an Associate Editor for IEEE OPEN JOURNAL OF ANTENNAS AND PROPAGATION, MEDICAL PHYSICS, AND DIAGNOSTICS.