Journal of
Sensor and Actuator Networks

MDPI

Article

Multi-Class Classification of Human Activity and Gait Events Using Heterogeneous Sensors

Tasmiyah Javed [1,2], Ali Raza [1,2,*], Hafiz Farhan Maqbool [1,3], Saqib Zafar [3,4], Juri Taborri [4] and Stefano Rossi [4,*]

1 Human-Centered Robotics Lab of NCRA, University of Engineering and Technology Lahore, Lahore 54890, Pakistan; tasmiyah.javed@gmail.com (T.J.); farhan.maqbool@uet.edu.pk (H.F.M.)
2 Department of Mechatronics and Control Engineering, University of Engineering and Technology Lahore, Lahore 54890, Pakistan
3 Department of Mechanical, Mechatronics and Manufacturing Engineering, University of Engineering and Technology Lahore, Faisalabad Campus, Faisalabad 38000, Pakistan; saqibzafar@uet.edu.pk
4 Department of Economics, Engineering, Society and Business Organization, University of Tuscia, 01100 Viterbo, Italy; juri.taborri@unitus.it
* Correspondence: aliraza@uet.edu.pk (A.R.); stefano.rossi@unitus.it (S.R.)

Abstract: The control of active prostheses and orthoses requires the precise classification of instantaneous human activity and the detection of specific events within each activity. Furthermore, such classification helps physiotherapists, orthopedists, and neurologists in kinetic/kinematic analyses of patients' gaits. To address this need, we propose an innovative deep neural network (DNN)-based approach with a two-step hyperparameter optimization scheme for classifying human activity and gait events, specific for different motor activities, by using the ENABL3S dataset. The proposed architecture sets the baseline accuracy to 93% with a single hidden layer and offers further improvement by adding more layers; however, the corresponding number of input neurons remains a crucial hyperparameter. Our two-step hyperparameter-tuning strategy is employed which first searches for an appropriate number of hidden layers and then carefully modulates the number of neurons within these layers using 10-fold cross-validation. This multi-class classifier significantly outperforms prior machine learning algorithms for both activity and gait event recognition. Notably, our proposed scheme achieves impressive accuracy rates of 98.1% and 99.96% for human activity and gait events per activity, respectively, potentially leading to significant advancements in prosthetic/orthotic controls, patient care, and rehabilitation programs' definition.

Keywords: convergent–divergent architecture; deep neural network (DNN); human activity classification (HAC); human activity recognition (HAR); gait event detection (GED)

check for updates

Citation: Javed, T.; Raza, A.; Maqbool, H.F.; Zafar, S.; Taborri, J.; Rossi, S. Multi-Class Classification of Human Activity and Gait Events Using Heterogeneous Sensors. *J. Sens. Actuator Netw.* **2024**, *13*, 85. https://doi.org/10.3390/jsan13060085

Academic Editor: Lei Shu

Received: 17 October 2024
Revised: 29 November 2024
Accepted: 4 December 2024
Published: 10 December 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The accurate and precise classification of human gait activity and gait events per activity has growing significance in the healthcare domain. It helps doctors/physiotherapists in the early detection of mobility issues, suggesting suitable rehabilitation programs, and monitoring a patient's recovery after interventions. Roboticists also need this classification for developing better and real-time control schemes for active prosthetic or orthotic devices [1]. Consider, for instance, the case of amputations that severely affect an individual's physical, psychological, and professional capacities, resulting in the degradation of one's quality of life (QOL) [1,2]. For any healthcare professional who handles such amputations and subsequent rehabilitation processes, it is vital to regularly monitor the quality of a patient's physical activities and assess the rehabilitation progress [3], making the classification of motor activities necessary. Amputees require prostheses to regain their mobility. Modern active prostheses must infer the intended locomotion profile from an integrated sensory

system and use actuators for imparting precisely controlled motion to artificial limb(s) in real-time [4]. Yet, the success of these devices hinges on two major objectives: human activity classification (HAC), and gait event detection (GED). Attaining both objectives with higher accuracy readily translates into improved prosthetic control. Similarly, powered orthoses require such a classification process.

Generally, two different approaches are adopted for the acquisition of motor data related to human activities. The first method employs an optoelectronic system such as Vicon, which utilizes infrared cameras and reflective markers to accurately capture and precisely track the movement dynamics in the three-dimensional motion of the body. The second method uses wearable sensors, such as accelerometers and gyroscopes, to gather musculoskeletal data by measuring joint angles and muscle activity [5]. Optoelectronic systems are typically limited to indoor environments, hence often falling short of adequately replicating outdoor/real-world environments [6]. Wearable sensors, on the other hand, can feasibly and reliably capture human walk data in scenarios that closely mimic routine activities, including indoors and outdoors [7]. Among wearable sensors, a plethora of different sensor systems can be found, including inertial measurement units (IMUs), goniometers, and surface electromyogram (sEMG) sensors [8]. Various earlier motion recognition studies are solely based on data acquired through either mechanical sensors such as accelerometers, gyroscopes, and force sensors, or surface electromyograms (sEMGs) to measure joint angles, muscle activity, and forces exerted during movement. However, heterogeneous sensors gather kinematic as well as kinetic information about an individual's motion patterns [9]. When put together, so-called heterogeneous sensory data, these pieces of information have the potential to reveal the range of human motion and its variations, and may also enhance the effectiveness of high-level controllers of advanced prosthetic devices [5].

Activity recognition, on the other hand, refers to identifying a person's specific mobility activity such as walking, sitting, standing, stair ascent/descent, ramp up/down, jumping, etc., whereas gait event detection is the identification of specific events/instances during an activity, such as Initial Contact (IC), Toe-Off (TO), heel contact (HC), etc. Gait event identification subsequently segments the gait cycle into gait phases and provides vital feedback information for controlling prosthetics and other wearable devices [10]. For example, Ref. [11] presents a general control scheme for prosthetic/assistive devices, in which the high-level controller manages intent recognition by integrating feedback signals from the prosthesis, the user, and the environment. In this context, feedback from the prosthesis refers to sensor data—such as joint angles or applied forces—that provide real-time information on the device's state, enabling precise adjustments to align with the user's intended movements, whereas the middle-level controller translates the user's intent to a desired state upon the detection of events/phases. Hence, the feedback information on gait events/phases can be linked to activity or intent recognition.

The human activity data, before being fed to the classifier, need signal processing [12]. This generally requires noise removal and the generation of statistical features which, afterward, are input to the classifier for training, validation, and testing [13]. Since the nature of the extracted features is diverse, many classification schemes with different degrees of complexity have been designed. These include simple threshold-based techniques [14], as well as advanced machine learning and deep learning algorithms [15–17].

On one hand, threshold-based classification has been successfully applied to different datasets to identify specific static postures including lying, sitting, and standing. Their data usually come from accelerometers mounted on the pelvis/trunk [18], lower limb segments [19], and chest [20]. In addition, postural transitions have also been identified from accelerometers or gyroscope data [20,21]. Unfortunately, this approach is unsuitable for the multi-class activity classification [14] of HAC and GED because it cannot handle a wider range of postures and activities. On the other hand, off-the-shelf machine learning techniques have been shown to largely overcome these limitations. Classification algorithms such as K-Nearest Neighbors (K-NN), Support Vector Machine (SVM), Naïve Bayes (NB), multi-layer perceptron (MLP), radial-based function (RBF), hybrid learning algorithms

(HLAs), extreme gradient boosting (XGB), and several shallow layered (with one hidden layer) artificial neural networks (ANNs) have been applied and tested on various datasets for classification [22–26]. Numerous studies show ANN is a strong pattern recognition technique, owing to its capability of splitting datasets by flexible non-linear classifiers [27]. However, since its implementations typically yield shallow networks (i.e., they only have a few hidden layers), it may not transmit enough information between atomic learning units (neurons) for classifying heterogeneous datasets such as those containing varying natural and disparate sensory signals [28]. Larger datasets of these types will increase the ANNs computational time and lower their accuracies [12].

DNNs increase the number of hidden layers of ANNs to enrich intermediate feature representations capable of unearthing otherwise hidden, more crucial discriminatory data attributes. With enhanced feature space, DNNs have more pathways to efficiently reach feasible/optimal non-linear classifiers for heterogeneous datasets [29]. In addition, DNNs reduce the processing time in comparison to the manual feature extraction. They automatically extract and learn a meaningful representation of features directly from the training data. These capabilities make them an appropriate choice for the multi-channel and multi-class classification of human locomotion.

The DNNs have shown promising results for various application domains including time series classification [30], text retrieval [31], and text document classification [32]. In addition, several studies [17,33–36] show the utilization of deep neural networks for the human activity classification of datasets as well as the fusion of data acquired from motion capture techniques and wearable sensors. A recent research work [37] implemented a DNN on a dataset collected from three motion capture sensors. This work achieved 95% accuracy but required larger memory because the method involves image processing. Moreover, the literature mentions the diverse architectures of DNNs such as Convolutional Neural Networks (CNNs) [36,38], Long-Short Term Memory (LSTM) networks, and Recurrent Neural Networks (RNNs) [19,21]. For integrating heterogeneous information and improving the classification performance of hyperspectral images (HSIs) and light detection and ranging data (LiDAR), an interleaving perception convolution neural network was proposed. To reconstruct the HSI and LiDAR data together, a bisectional encoder was designed in which the reconstruction process is trained and then passed through a two-branch CNN for final classification without any dependency upon the annotated information [27]. An efficient hierarchical random walk network (HRWN) was proposed for the investigation of HIS and LiDAR data. A dual tunnel CNN was developed to capture spectral and spatial features, capturing the contrast between them by a pixel-wise affinity branch [28]. A hyperspectral anomaly detection method based on fractional Fourier entropy (FrFE) was used for preprocessing and optimal fractional transform order determination. To discriminate the anomalies from the background noise, FrFE obtains intermediate domain features between the original reflectance spectrum and its Fourier transform by space–frequency representation [39].

However, to the best of our knowledge, the ENABL3S dataset [8] has not been used for this purpose, nor has an optimal DNN been designed for this dataset. Compared to datasets such as Pamap2 [40] or Opportunity [41], ENABL3S offers a broader scope of activities, including rare scenarios like ramp ascent/descent and stair navigation. Additionally, ENABL3S integrates data from diverse sensors, providing a richer feature set that enables more precise activity and gait event classification. While many datasets focus exclusively on IMU or EMG data, combining these and goniometer data in ENABL3S captures both kinematic and kinetic information, making it particularly valuable for heterogeneous sensor analysis. In this context, and in conjunction with Figure 1, salient features of our work are summarized below:

1. We present a general purpose multi-class classification scheme for human activities using a new convergent–divergent DNN. The scheme successfully classifies several human activities such as level ground walk, stair ascent/descent, and ramp

ascent/descent among other classes, and within an activity, different gait events such as heel contact, pushing, toe-off, and free swing.

2. The DNN architecture, named Deep Multi-class Human Activity Classifier (DeMul-HAC), interprets large heterogeneous multi-sensor data and exploits the discriminating power of the ENABL3S complementary feature set using the compact statistical metrics (e.g., mean, variance, max, and min. values) of segmented sensor signals.
3. Inspired by heuristic-based bi-directional random search methods, which utilizes heuristic strategies (rules of thumb) to find the solution space from both the starting and target points for the complex problem [42], we propose a novel two-step hyper-parameter optimization scheme using a convergent–divergent architecture. It first searches for an appropriate number of hidden layers, using initially a convergent and later a convergent–divergent exploration scheme, and then carefully selects the number of neurons across these layers using 10-fold cross-validation.
4. Our proposed scheme elevates the test classification accuracy for human activity and gait events per activity to an impressive 95.4% and 99.83%, respectively, within the optimal processing time.
5. The designed classifier significantly outperforms other state-of-the-art methods such as decision trees, SVM, K-NN, etc., used for multi-class classification in the literature.

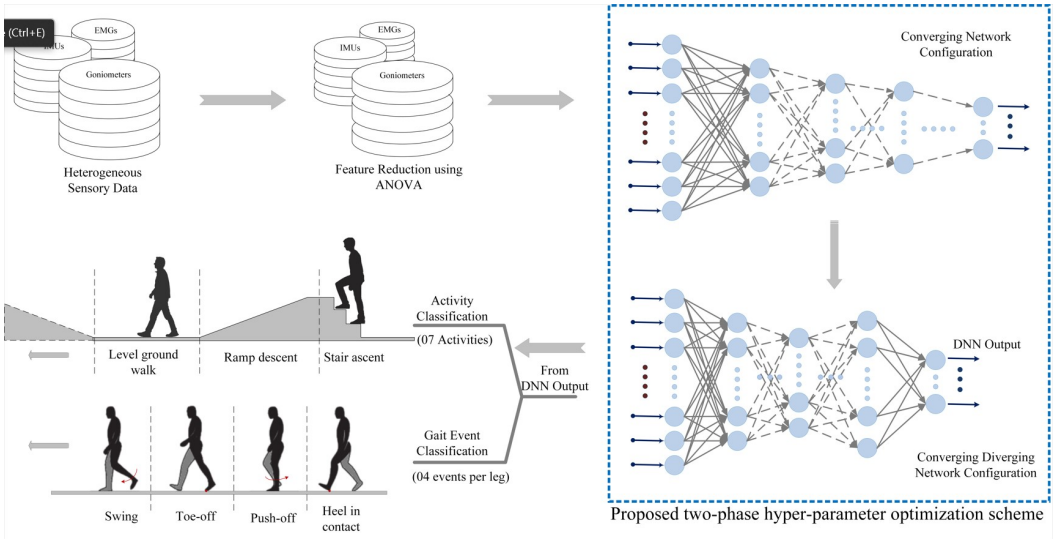


Figure 1. An overview of the proposed framework of the DNN-based approach, highlighting the two-phase hyperparameter optimization scheme and its application on multi-activity and multi-gait event classification.

The initial hypothesis of the study was that a deep neural network (DNN) optimized through a convergent–divergent architecture and fed with data from multiple heterogeneous sensors (such as IMUs, sEMGs, and goniometers) would enable the high-precision classification of human activities and the detection of gait events. This hypothesis was based on the assumption that the integration of varied sensor data could provide comprehensive, discriminative information across different activities and gait phases, which traditional machine learning methods might not capture effectively.

The remaining discussion is structured as follows. Section 2 describes the proposed methodology together with the dimensionality reduction method and the two-phase optimization algorithm for the DNN architecture selection. Section 3 discusses the experimental results and compares them with the results from existing state-of-the-art methods. At the end, Section 4 sums up the conclusions and points to future research directions.

2. Related Work

Several studies have explored various classifiers and datasets to enhance activity and gesture recognition, each contributing unique insights and methodologies. Rebelo et al. [43] demonstrated the effectiveness of combining accelerometer, EMG, and goniometer data for human activity recognition using a Hidden Markov Model (HMM). Their approach achieved an accuracy of 79%, highlighting the advantages of integrating multiple sensor modalities to improve continuous activity recognition and event detection. This multi-sensor combination provided a more comprehensive understanding of complex movements by leveraging the strengths of each sensor type. Hammerla et al. [44] investigated deep learning models, including DNN, CNN, and LSTM, on datasets such as PAMAP2, Opportunity, and Daphnet Gait. Their work emphasized the impact of learning rate, regularization, and architecture on model performance, evaluated using the F-measure. However, their findings were limited to offline implementations, restricting the real-time applicability of their approach. San et al. [45] applied CNNs to the OAR, REALDISP, and Hand Gesture datasets, achieving accuracies of 86.7%, 96%, and 90.1%, respectively. Their research focused on optimizing high-performing architectures tailored to these specific datasets, but it did not explore how these models might generalize across a broader range of activities or real-world settings. Xu et al. [46] proposed a two-branch CNN for HSI-LiDAR data, achieving a high accuracy of 99.13%. Similarly, Zhang et al. [47] employed PToP CNN on the same data type, reaching 89% accuracy. Both studies concentrated on architecture optimization for improved performance, but their focus on a specific data type may limit the broader applicability of their methods to more general activity recognition tasks. Ignatov et al. [48] used CNNs on the WISDM, Smartphone, and UCI datasets, achieving an accuracy of 94.35%. Their approach introduced variable window lengths for real-time feature extraction and explored hyperparameter variations to optimize performance. However, the reliance on specific datasets may affect the generalizability of their findings to other real-world scenarios. Moya Rueda et al. [49] evaluated CNN performance on Opportunity, PAMAP2, and order-picking datasets, achieving an accuracy of 92.22%. While they assessed different architectures, their work primarily focused on learning rate adjustments and did not explore other crucial aspects like multi-modal sensor fusion or real-time processing. Hassan et al. [50] employed a Deep Belief Network (DBN) with accelerometer and gyroscope data, achieving an impressive 97.5% accuracy. Their work involved hyperparameter tuning to design the best-performing network but was limited to offline implementations, thereby restricting its use in real-time applications. A comparison of the proposed study with the existing related work is also shown in Table 1.

Despite the advancements achieved by these studies, a few common limitations persist across many approaches. First, many studies primarily rely on a single sensor type or less diverse sensor datasets, which may fail to capture the full range of human motion and context. Additionally, most approaches are limited to offline analysis, restricting their applicability in real-time settings where continuous, accurate detection is crucial. Furthermore, while many studies focus on optimizing model architectures, they often neglect the importance of real-time, adaptive processing and sensor fusion across different modalities for comprehensive activity recognition.

The novelty of this research lies in the introduction of a two-phase optimization strategy for a deep neural network (DNN) designed for the multi-class classification of human activities and gait events. This work is motivated by the growing need for precise classification in applications such as prosthetic and orthotic control, where the real-time and accurate detection of human motion is critical. By leveraging the ENABL3S dataset, which integrates heterogeneous sensor data, including IMU, sEMG, and goniometer signals, the proposed framework addresses the limitations of previous approaches that rely on single-sensor or less diverse datasets, offering a more comprehensive solution for human activity recognition.

Table 1. Comparison of proposed study with existing related work.

Ref.	Dataset	Classifier	Accuracy	Capabilities	Limitations
Rebelo et al. [43]	Accelerometer EMG Goniometer	HMM	79%	Effect of combinations of sensor data	Continuous recognition of activities Events recognition
Hammerla et al. [44]	Pamap2 Opportunity Daphnet Gait	DNN CNN LSTM	F-measure	Model performance w.r.t. learning rate, regularization & architecture	Offline implementation
San et al. [45]	OAR Hand gesture REALDISP	CNN	86.7% 96% 90.1%	Design of best performing architecture	Search for architectures
Xu et al. [46]	HSI LiDAR	Two Branch CNN	99.13%	Data Classification	Search for architectures
Zhang et al. [47]	HSI LiDAR	PToP CNN	89%	Feature Extraction and Classification	Search for architectures
Ignatov [48]	WISDM Smartphone UCI	CNN	94.35%	Variable window length for feature extraction Real-time	Search for architectures and variation of hyperparameters
Moya Rueda et al. [49]	Opportunity Pamap2 Order picking data	CNN	92.22%	Performance evaluation with different architectures & datasets	Only learning rate is studied
Hassan et al. [50]	Accelerometer Gyroscope	DBN	97.5%	Tuning of hyperparameters for best performing network	Offline implementation
Proposed Study	Accelerometer EMG Goinometer	DNN	98.1% 99.96%	activity recognition Gait Event Detection	Offline implementation

3. Methodology

This section details the proposed methodology for human activity classification (HAC) and gait event detection (GED) of multi-channel and multi-class data, using heterogeneous sensors such as accelerometer, goniometer, and sEMG. The initial steps in classification, after the data acquisition, include data analysis and preprocessing. Then, deep neural network architectures are modeled and evaluated with the help of standard figures of merit. The study also investigates the properties of heterogeneous sensor signals and ranks the sensors according to the importance of feature space used for training the DNN model. Readers are also referred to the schematic diagram (Figure 1) of the proposed framework.

3.1. Data Description

The Encyclopedia of Able-Bodied Bilateral Lower Limb Locomotor Signals (ENABL3S) is a publicly available dataset. It comprises time series data of lower limb activity acquired through sEMG, IMU, and goniometer sensors [8]. The data contain the action potential of seven muscle locations from each leg, i.e., tibialis anterior (TA), soleus (SOL), medial gastrocnemius (MG), biceps femoris (BF), vastus lateralis (VL), rectus femoris (RF), and semitendinosus collected through surface electromyograms (sEMG). The data also consist of the angular acceleration and the velocity of each segment of the lower limbs. Specifically, 6 DOF inertial measurement units (IMUs) are bilaterally attached to the shank, thigh, and waist. Moreover, the angular displacement of the knee and ankle joints is also recorded by placing goniometers on each joint, respectively.

The dataset, denoted as ***EN*** comprises the multi-channel and multi-class data of seven different walking activities and four different gait events acquired from ten healthy subjects (seven male, three female; 25.5 ± 2 years; 174 ± 12 cm; 70 ± 14 kg) with no gait impairments. The total repetitions for each activity are 25. The set of activities (level ground walking, stair ascent, stair descent, ramp ascent, ramp descent, standing, and sitting) is referred to as $\boldsymbol{A}$, whereas the set of gait events (right heel contact, right toe-off, left heel contact, and left toe-off) is represented as $\boldsymbol{G}$. The total number of participants is denoted by $\boldsymbol{P}$, whereas t represents the trial number of the p^{th} participant of the experiment. The feature vector is represented with the symbol $\boldsymbol{D}$ with dimensions of the feature set denoted by $\boldsymbol{F}$, where $\boldsymbol{f}$ is the indices, i.e., $\boldsymbol{f} = 1, 2, 3, \ldots,$ F for any positive integer F. Similarly, $D = \sum_1^p D_p$ is the feature vector evaluated from the data of $\boldsymbol{P}$ participants. Using the notation ***D(P,t)*** for the collective feature vector calculated from the data of t trials, we have $D(P,t)^{(\mathrm{F})} = d_{P,t}{}^{(1)}, d_{P,t}{}^{(2)}, d_{P,t}{}^{(3)}, \ldots, d_{P,t}{}^{(\mathrm{F})}$, where superscripts are indicating the index of the calculated features from the acquired sensor data.

3.2. Dimensionality Reduction

Higher-dimensional datasets not only require bigger memory space but also increase processing time for the classification [51]. We have applied dimension reduction techniques [52] to remove redundant and noisy data to overcome the risks of overfitting. Feature reduction aims to present N-dimensional data:

$$\overline{E_N} = \overline{D^F_{P,t}} = \overline{d^{(F)}_{p,t}} : \overline{d^{(f)}_{p,t}} \in R^N \tag{1}$$

by M-dimensional data space as

$$\overline{E_N} = \overline{D^F_{P,t}} = \overline{d^{(F)}_{p,t}} : \overline{d^{(f)}_{p,t}} \in R^M \tag{2}$$

where M < N.

This research employs multi-factor ANOVA (Analysis of Variance) to create a lower-dimensional feature space from the feature set $(D_{P,t}{}^{(\mathrm{F})})$. Multi-factor ANOVA is selected for its ability to reduce cost, complexity, and processing time compared to PCA and LDA. Unlike these global techniques, ANOVA provides insights into the impact of individual variables and their interactions by analyzing p-values ($\leq$0.05) derived from the comparison

of response variable means across feature combinations [53]. The threshold in this research is set to p-value ≤ 0.05. This implies that the features under this criterion greatly affect the response variables, as they have high variability.

Moreover, another feature reduction technique, Minimum Redundancy Maximum Relevance (mRMR) is applied to this dataset. mRMR computes mutual information between the feature set and output classes. Classification variables (features) are selected based on mutual dissimilarity among themselves and marginal similarity to the response variables [54]. This approach calculates the Maximal Relevance (*Rmax*) in the first stage and Minimum Redundancy (*Rmin*) in the second stage, respectively. The value of mutual information is determined in the final step of mRMR, after the calculation of both (*Rmax*) and (*Rmin*).

The ENABL3S data have 368 features, which are reduced to 260 features by using multi-factor ANOVA and 42 features with mRMR. Both of these feature reduction methods are applied independently for the modeling of the classifier. Time domain features such as mean, standard deviation, minimum, maximum, initial, and final values are considered for IMU and goniometer. For sEMG signals, the mean absolute value, number of zero crossings, slope sign changes, waveform length, and autoregressive coefficients of the sixth order are considered features. In total, there are 368 features, where 180 belong to IMU, 140 belong to sEMG, and 48 belong to the goniometer. The performance parameters for measuring the accuracy of the model selected in this research are mean squared error (MSE), confusion matrix, and comparison of model performance with other state-of-the-art machine learning techniques.

4. DNN Architecture and Modeling

Since musculoskeletal signals exhibit non-linear properties, mapping these input signals to their outputs using a mathematical model is difficult [55]. However, neural networks are significant tools able to efficiently map non-linear inputs to their outputs. Hence, neural networks are being deployed for the learning and classification of such data. Similarly, deep architectures of neural networks are also valuable for processing multiple data arrays in speech signals, image data, or sensor-based signals [56]. Therefore, this paper proposes a deep feed-forward neural network architecture for the classification of activities and gait events that occur as these activities are being performed.

The convergent–divergent architecture is the central proposition of the presented research. This term highlights the two-phase process of the neural network configuration. The first phase explores the appropriate number of hidden layers in an iterative fashion, where each iteration increments a hidden layer in the previous network but with a decreasing number of neurons in it, attempting to maximize the classification accuracy. Since the number of neurons is reduced in successive layers as illustrated in Figure 1, the term convergent is used. Once the classification accuracy starts to stagnate (minima), the second phase ensues, where a convergent–divergent approach is employed, in the vicinity of the selected network of the first phase, to search for the appropriate number of neurons in each layer to further enhance the classification accuracy.

4.1. Phase 1 Converging Networks

The first phase performs a sequential line search on 10 different DNN architectures starting from a single hidden layer to 10 hidden layered networks (DNNs). The in-line architecture is a converging network configuration, i.e., the neuron number in hidden layers converges progressively with each increasing hidden layer as illustrated in Figure 1 as well.

Neurons in the first hidden layer are selected according to Equation (3) with 10% tolerance, while the neurons in the upcoming hidden layer are decreased by 10 neurons in each iteration:

$$N_n = 2\sqrt{(Y+2)X} \tag{3}$$

where X is the number of inputs, and Y is the number of outputs [57].

Algorithm 1 describes the process behind the converging network strategy. In each iteration of the algorithm, a network is formed, having a particular number of neurons and hidden layers. The networks are 10-fold cross-validated to search for the optimal architecture that achieves the highest intent recognition accuracy. Then, the average accuracy values corresponding to each network are calculated and compared. Networks with the highest average accuracy are considered starting suggestions for the second stage of the proposed strategy.

Algorithm 1 : Phase 1 of the proposed two-phase optimization strategy.

Require: Reduced feature set $\overline{D^{F}_{(P,t)}}$, DNN parameters Θ ▷ Input parameters
for $i = 1$ to m **do** ▷ Averaging of results
  for $j = 1$ to l **do**
    for $k = 1$ to 10 **do**
      Fit the DNN model on updated training set
      Validate the DNN model on updated validation set ▷ 10-fold cross validation
    end for
    Calculate average accuracy on validation set
  end for
  Test the DNN model
end for
Calculate and compare average MSE and accuracy of all DNN models
Arrange network models in descending order with respect to MSE and accuracy
return Performance metrics array in descending order

4.2. Phase 2 Converging–Diverging Networks

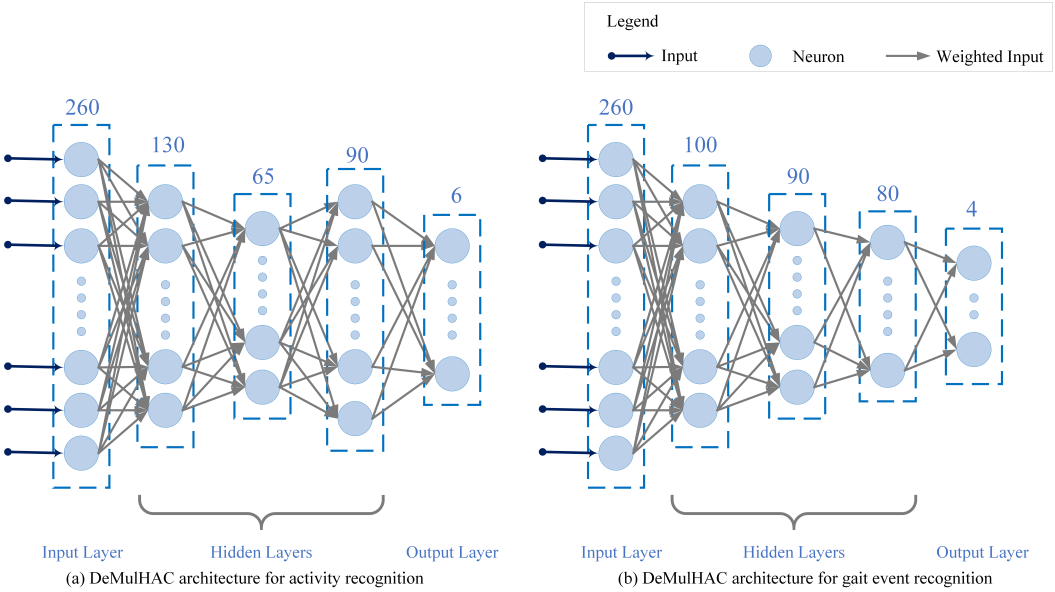
The second phase of DNN modeling adopts a hybrid form of commonly used grid search and random search methods [39]. Unlike conventional grid search, which can be computationally expensive due to exhaustive parameter space exploration, or random search, which risks missing critical parameter configurations, the proposed two-phase strategy systematically combines the strengths of both approaches. By initially narrowing the parameter space through a convergent approach, the method reduces the computational overhead. Subsequently, the diverging stage fine-tunes the architecture, enabling higher accuracy while maintaining efficiency. This hierarchical strategy ensures robust generalization across heterogeneous datasets like ENABL3S, which exhibit non-linear relationships and high dimensionality. The algorithm is designed based on the suggested performance results of phase 1. It is seen that increasing the hidden layers improves the accuracy values. It is also observed that the accuracy value saturates 6 hidden layers onward, while the total processing time keeps increasing with the increasing hidden layers. Networks with 4, 5, and 6 hidden layers produce maximum accuracy but have a relatively high processing time. Therefore, phase 2 is designed for further optimization with a converging–diverging configuration of hidden layers. Considering the results of phase 1, this phase is initialized from two hidden layers. The search is terminated on four hidden layers because the results started depicting saturation. The number of neurons in each hidden layer is defined using a random search method because of its better results than the grid search [39]. However, a grid search is performed for the final selection of the best-performing network.

Algorithm 2 describes the steps followed in this stage. The number of layers and neurons is varied according to the user input for the formation of a network in each iteration. Furthermore, each network is 10-fold cross-validated for verification of the best performing network. Figure 2a presents one such configuration for activity classification.

Algorithm 2 : Phase 2 of proposed two-phase optimization strategy.

Require: Reduced feature set $\overline{D^{F}_{(P,t)}}$, DNN parameters Θ

Enter the number of hidden layers l ▷ Input parameters
for $i = 1$ to m **do** ▷ Averaging of results for induced hyperparameters
  for $j = 1$ to l **do**
    Enter the number of neurons n in each layer ▷ Enumeration of neurons in each iteration
    for $k = 1$ to 10 **do**
      Fit the DNN model on updated training set
      Validate the DNN model on updated validation set ▷ 10-fold cross validation
    end for
    Calculate average accuracy on validation set
  end for
  Test the DNN model
end for
Calculate and compare average MSE and accuracy of all DNN models
Arrange network models in descending order with respect to MSE and accuracy
return DeMulHAC {DNN with performance}



(a) DeMulHAC architecture for activity recognition (b) DeMulHAC architecture for gait event recognition

Figure 2. Three hidden layered DeMulHAC architectures for activity and gait event classification.

A similar two-stage strategy is applied for the classification of gait events. The inputs of the system are the same 260 features $\overline{D_{P,t}^{F}}$ that are used for activity classification, while the outputs of the network are four gait events (right heel contact (RHC), right toe-off (RTO), left heel contact (LHC), and left toe-off (LTO)):

$$\text{Minimize } Z(b,\omega) = \frac{1}{2}\sum_{x=1}^{X}\|y_x - \hat{y}_x\|^2 \tag{4}$$

In each run (a single run constitutes the training, validation, and testing of the DNN model), authors randomly select 80% of the data samples for training and validation; the remaining 20% is reserved for testing. For a particular combination of hyperparameter values, a 10-fold cross-validation scheme is applied on the 80% chunk by dividing it into 10 folds of equal size. Nine out of ten folds are reserved for training, and the remaining fold is used as validation data. To ensure robust generalization and mitigate overfitting, we implement a rigorous 10-fold cross-validation approach, validating the model across diverse data partitions [58]. Dimensionality reduction techniques, such as multi-factor

ANOVA and mRMR, are utilized to eliminate noise and redundancy in the feature set. Additionally, the convergent–divergent scheme outlined in Algorithm 2 further reduces model complexity by optimizing layers and neuron counts. These combined measures enhance the model's performance, ensuring reliability across varied scenarios [59]. This splitting is performed 10 times, each time using a separate single fold as validation data. The training and validation accuracy is recorded in each iteration of the 10-fold cross-validation procedure as well as the average of the 10 readings. The whole procedure is repeated for a different combination of hyperparameter values, and the values of average training and validation accuracy are recorded. Network training is performed with the backward propagation of errors. Since the errors to be calculated for the entire dataset are a summation of individual errors, minimization of the objective function (Z) according to Equation (4) is performed, where X reflects the total number of inputs in the form of a $260X1$ vector, and y is the actual output, while $\hat{y_x}$ is the predicted output of a sample. To minimize this objective function, the Scaled Conjugate Gradient (SCG) algorithm is adopted for the adjustment of weights and biases in the direction of the gradient [60]. The developed model does not suffer from overfitting on the training data and performs reasonably well on the test data. Hence, we do not use any weight regularization strategy such as L1, L2 regularization, or dropout regularization. The DNN modeling, learning, and testing are carried out with the MATLAB 2018a toolbox installed on an DELL i7-4790 CPU, 360 GHz processor with 8 GB RAM. The following Table 2 shows the detailed parameters for the different layers of the DeMulHAC architecture as per the algorithms described above.

Table 2. Detailed parameters for the different layers of the DeMulHAC architecture.

Parameter	Human Activity Classification (HAC)	Gait Event Detection (GED)
Neurons of Input Layer	260	260
Neurons of Output Layer	6	4
Neurons of Hidden Layers	130, 65, 90	100, 90, 80
Activation Function: Hidden Layers	Hyperbolic tan (tanh)	Hyperbolic tan (tanh)
Activation Function: Output Layer	Softmax	Softmax
Optimization Algorithm	Scaled Conjugate Gradient	Scaled Conjugate Gradient

The selection of parameters is conducted upon technical considerations as follows. The number of neurons in each hidden layer is optimized using a two-phase strategy: Phase 1 systematically reduces neurons across layers to enable hierarchical feature learning, while Phase 2 employs a converging–diverging approach to fine-tune neuron counts, enhancing generalization for heterogeneous sensor data. The tanh activation function is chosen for hidden layers due to its stability with signals of mixed polarity, and softmax is applied in the output layer for probabilistic multi-class classification. The Scaled Conjugate Gradient (SCG) optimizer, which does not require manual learning rate tuning, is used to efficiently manage the high-dimensional feature space [61–64].

5. Results and Discussions

Neural network modeling includes choosing an optimal network topology, training algorithm, and architecture to achieve the desired performance. To enhance the conventional intent recognition studies, we propose a two-phase optimization strategy based on DNN and ENABL3S data. Performance parameters for measuring the accuracy of the model selected in this research are the mean squared error (MSE), confusion matrix, and comparison of model performance with other state-of the-art machine learning techniques.

5.1. Activity Classification

Tables 3 and 4 present the comparison results of all networks in Phase 1 and Phase 2 using the proposed two-phase optimization strategy adopted for activity classification. Table 5 depicts a decrease in MSE and an increase in accuracy with the increasing number of hidden layers until 6 hidden layers, which seems to be a saturation point. However,

higher accuracy (comparable to network 6) is again achieved with 10 hidden layered networks. Despite this, a significant rise in the total processing time is observed for the deeper network. The processing time for the model is evaluated for different network depths. For the shallowest network (1 hidden layer), the average processing time per sample is 10 ms, while for deeper networks (up to 6 hidden layers), the average time increases to 50 ms per sample. Despite the increase in processing time, accuracy continues to improve, with the optimal trade-off observed at four hidden layers, where the system achieves 99.4% accuracy for activity classification while maintaining a processing time of 30 ms per sample. These benchmarks highlight the trade-off between model depth and computational cost, with deeper networks offering diminishing returns in terms of accuracy improvement relative to the increase in processing time. As we are targeting a considerably high-performing network, having an optimal processing time, the observations lead to the next stage of network architecture selection, having the converging–diverging configuration of networks.

Table 3. Phase 1: comparison of converging networks for activity classification.

Network	Hidden Layer Neurons	MSE	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)	Overall Accuracy (%)
1	100	0.00112 ± 0.0015	96.1 ± 0.89	94.5 ± 1.64	93.8 ± 0.92	94.8 ± 0.74
2	100, 90	0.000711 ± 0.0036	95.3 ± 1.23	95.6 ± 1.06	94.2 ± 0.95	95.0 ± 1.17
3	100, 90, 80	0.00108 ± 0.0047	96.7 ± 0.93	95.2 ± 0.73	94.2 ± 0.85	95.4 ± 0.90
4	100, 90, 80, 70	0.000893 ± 0.0006	97.9 ± 0.61	95.5 ± 0.71	95.3 ± 0.54	97.1 ± 0.63
5	100, 90, 80, 70, 60	0.00095 ± 0.0007	99.4 ± 0.68	94.9 ± 0.73	95.8 ± 0.63	98.2 ± 0.70
6	100, 90, 80, 70, 60, 50	0.00126 ± 0.0034	99.1 ± 0.94	95.3 ± 0.92	94.3 ± 0.89	97.8 ± 0.90
7	100, 90, 80, 70, 60, 50, 40	0.000753 ± 0.0083	98.9 ± 1.32	94.7 ± 1.03	94.4 ± 1.52	97.6 ± 1.42
8	100, 90, 80, 70, 60, 50, 40, 30	0.000165 ± 0.0072	97.7 ± 1.00	94.1 ± 1.45	94.1 ± 1.36	96.6 ± 1.18
9	100, 90, 80, 70, 60, 50, 40, 30, 20	$1.20 \times 10^{-11} \pm 0.0003$	98.0 ± 1.03	94.1 ± 1.36	93.9 ± 1.54	96.8 ± 1.90
10	100, 90, 80, 70, 60, 50, 40, 30, 20, 10	$8.25 \times 10^{-9} \pm 0.0007$	99.4 ± 1.93	95.1 ± 1.98	93.5 ± 2.82	97.9 ± 2.63

Table 4. Phase 2: comparison of converging–diverging networks for activity classification.

Network	Neurons in Hidden Layers	MSE	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)	Overall Accuracy (%)
2	130, 65	0.0082 ± 0.0037	98.08 ± 1.92	94.60 ± 2.03	94.53 ± 1.70	97.04 ± 1.86
2	100, 50	0.0079 ± 0.0036	98.19 ± 1.98	94.68 ± 1.51	94.37 ± 1.95	97.07 ± 1.81
2	160, 80	$0.0064 \pm 9.52 \times 10^{-4}$	98.97 ± 0.36	95.17 ± 0.49	95.08 ± 0.26	97.82 ± 0.31
3	130, 65, 90	0.0008 ± 0.0011	99.30 ± 0.45	94.53 ± 0.47	95.40 ± 0.40	98.10 ± 0.37
3	100, 50, 75	0.0094 ± 0.0034	97.43 ± 1.74	95.50 ± 1.92	94.26 ± 1.72	96.51 ± 1.75
3	160, 80, 105	0.0085 ± 0.0023	97.98 ± 1.00	94.87 ± 1.35	94.60 ± 0.99	96.99 ± 1.03
4	130, 65, 90, 45	0.0073 ± 0.0011	98.44 ± 0.50	95.05 ± 0.61	94.83 ± 0.22	97.39 ± 0.38
4	100, 50, 75, 38	$0.0083 \pm 9.85 \times 10^{-4}$	98.03 ± 0.44	95.14 ± 0.75	94.89 ± 0.19	97.13 ± 0.36
4	160, 80, 10, 53	0.0081 ± 0.0020	98.10 ± 1.00	94.90 ± 1.05	94.38 ± 0.93	97.04 ± 0.97

Table 5. Comparison of converging networks for gait event classification.

Network	Neurons in Hidden Layers	MSE	Training Accuracy	Validation Accuracy	Testing Accuracy	Overall Accuracy
1	100	0.0021 ± 0.0013	97.3 ± 0.98	94.5 ± 1.64	96.8 ± 0.89	97.0 ± 0.90
2	100, 90	0.0015 ± 0.0010	98.6 ± 0.91	95.6 ± 1.06	98.2 ± 1.12	98.4 ± 0.98
3	100, 90, 80	0.00079 ± 0.0005	99.8 ± 0.00	95.2 ± 0.73	99.83 ± 0.11	99.96 ± 0.05
4	100, 90, 80, 70	0.0032 ± 0.0011	98.9 ± 0.83	95.5 ± 0.71	98.3 ± 0.87	98.8 ± 0.85
5	100, 90, 80, 70, 60	0.0043 ± 0.0023	98.4 ± 1.00	94.9 ± 0.73	97.2 ± 1.60	97.7 ± 1.54
6	100, 90, 80, 70, 60, 50	0.0042 ± 0.0020	98.1 ± 0.54	95.3 ± 0.92	97.8 ± 0.62	98.1 ± 0.47
7	100, 90, 80, 70, 60, 50, 40	0.0031 ± 0.0011	97.9 ± 0.56	94.7 ± 1.03	97.4 ± 0.59	97.8 ± 0.57
8	100, 90, 80, 70, 60, 50, 40, 30	0.0033 ± 0.0016	97.7 ± 0.59	94.1 ± 1.45	96.2 ± 0.68	96.9 ± 0.63
9	100, 90, 80, 70, 60, 50, 40, 30, 20	0.0090 ± 0.0038	98.0 ± 1.21	94.1 ± 1.36	97.1 ± 1.34	97.6 ± 1.15
10	100, 90, 80, 70, 60, 50, 40, 30, 20, 10	0.0082 ± 0.0029	99.4 ± 1.04	95.1 ± 1.98	99.83 ± 0.97	99.96 ± 0.92

The second phase involves testing 2, 3, and 4 hidden layers with different sets of neurons. A similar trend to the previous phase is observed in the second phase as well. Accuracy percentage saturates with the increasing depth of the network with increased processing time. Hence, the hidden layer number is not incremented beyond four in converging–diverging architectures. The explained trend, in conjunction with the number of neurons, as exhibited by the network architecture, is displayed in Figure 3.

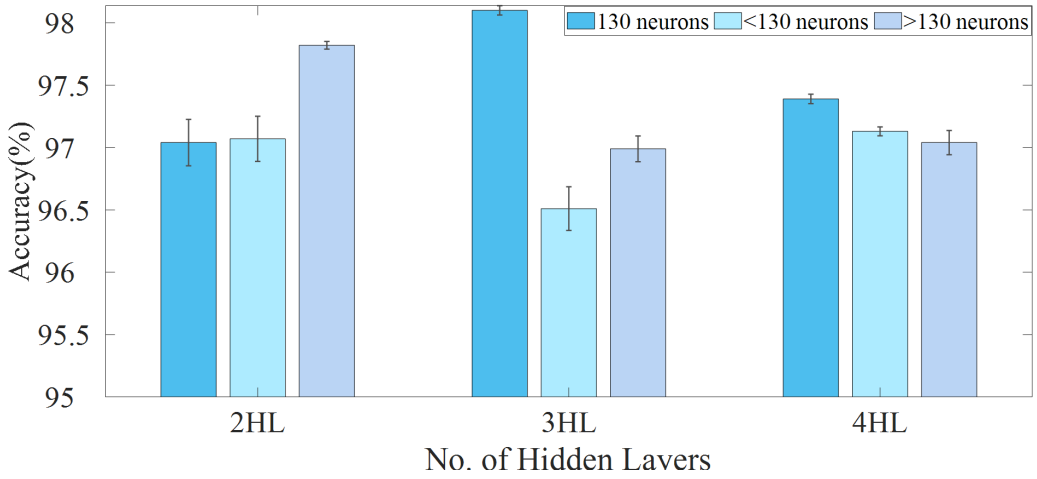


Figure 3. DNN accuracy for activity recognition in phase 2.

Finally, the outperforming network among all the tested networks is selected. As mentioned earlier, this network is named a Deep Multi-class Human Activity Classifier (DeMulHAC) and is shared as open-source data, ensuring transparency and reproducibility. The corresponding confusion matrix for activity recognition is shown in Table 6. The diagonal elements represent the number of correctly classified samples, while the non-diagonal values show the number of times a class has been misclassified. Accuracy for standing activity is slightly decreased because the data for this activity are comparatively limited in the available dataset. But other than that, the achieved accuracy of DeMulHAC for activity classification is 98.1 ± 0.37%.

Table 6. Confusion matrix for activity classification with precision, recall, F1-Score, and error rate.

Target Class	LW	RA	RD	SA	SD	St	Precision (%)	Recall (%)	F1-Score (%)	Error (%)
Output Class										
LW	11,595	26	67	13	20	78	98.3	98.4	98.3	1.7
RA	28	3207	0	6	0	0	99.0	98.3	98.7	1.4
RD	50	0	3838	0	3	2	98.6	98.3	98.5	1.0
SA	28	3	0	1398	1	0	97.8	98.6	98.2	2.2
SD	30	9	0	1	1382	1	97.1	98.2	97.6	2.9
St	53	16	1	0	2	533	88.5	87.2	87.8	11.5
Recall (%)	98.4	98.3	98.3	98.6	98.2	87.2				
Error (%)	1.6	1.7	1.7	1.4	1.8	12.8				

5.2. Gait Event Detection

Table 7 displays all the networks and their corresponding accuracies for gait event classification.

For both classification tasks, the network with three hidden layers shows the highest performance, outperforming all other network configurations in the first phase of the adopted strategy. Therefore, the search is terminated only after the first optimization stage. Figure 4 displays the results obtained in this phase.

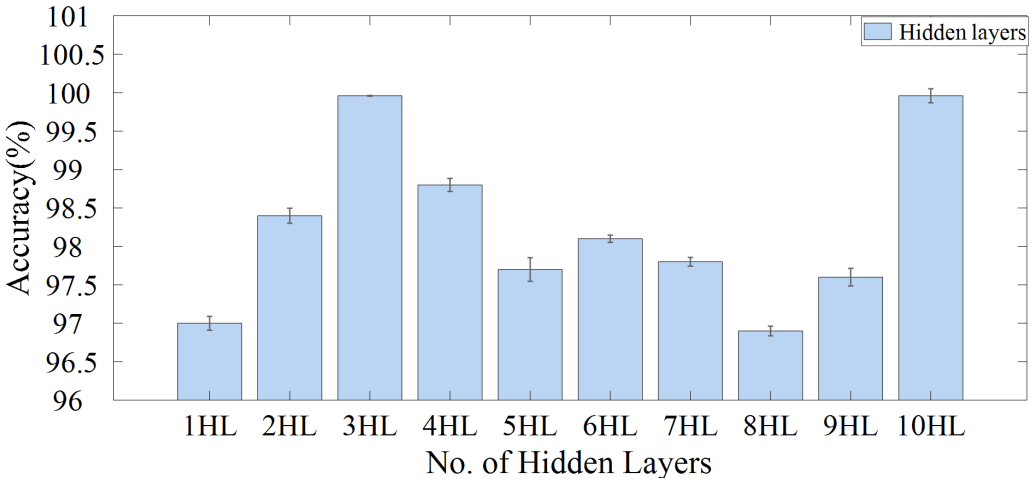


Figure 4. DNN accuracy for gait event detection.

The DeMulHAC network achieves an accuracy of 99.96 ± 0.05% for gait event classification. All gait events are accurately classified with the fewest misclassifications as shown in Table 7. The DeMulHAC architectures for both activity and gait event classification are displayed in Figure 2.

Table 7. Confusion matrix for gait event detection with precision, recall, F1-Score, and error rate.

Target Class	RHC	RTO	LHC	LTO	Precision (%)	Recall (%)	F1-Score (%)	Error (%)
Output Class								
RHC	5401	1	1	0	100	99.7	99.9	0
RTO	0	5711	0	3	99.9	99.9	99.9	0.1
LHC	18	0	5697	1	99.7	99.9	99.8	0.3
LTO	0	3	2	5773	99.9	99.9	99.9	0.1
Recall (%)	99.7	99.9	99.9	99.9				
Error (%)	0.3	0.1	0.1	0.1				

The proposed model demonstrates high precision and recall (Table 6) across most lower limb activities, indicating strong performance in both correctly identifying activities and minimizing false positives. The F1-Score (Table 6), reflecting a balance between precision and recall, is generally high, though the "St" (Standing) activity shows slightly lower performance, with an elevated error rate of 11.5%. This suggests that the "St" activity is more challenging for the model to classify correctly, contributing to its higher error rate compared to other activities, while the overall error rate remains low for most activity categories. In standing activity, the body's movement is minimal and has no specific movement pattern, thus making it more susceptible to noise or likely to lead to misclassification as compared to other activities because other activities have distinct patterns for classification. The proposed model also demonstrates near-perfect precision and recall (Table 7), highlighting its exceptional accuracy in detecting each gait event. The F1-Score (Table 7) is consistently high across all events, reflecting well-balanced performance in both precision and recall. The error rate is extremely low for all gait events, with the highest error being just 0.3% for the "LHC" event, which is still an excellent result. This indicates that the model performs robustly, making accurate predictions with minimal misclassifications.

5.3. Comparison with State-of-the-Art Methods

The DeMulHAC network is compared with the Support Vector Machine (SVM), decision tree (DT), K-Nearest Neighbors (kNN), and the ensemble method of boosted trees for accuracy and processing time to provide a well-rounded evaluation of our approach against a set of widely used, robust algorithms in the field of classification. These methods are standard benchmarks in human activity recognition (HAR), gait event detection, and similar classification tasks, and they offer diverse approaches in their learning mechanics, which is essential for demonstrating the broad applicability and strengths of DeMulHAC [65]. Figure 5 clearly illustrates that for ENABLE3S activity classification, the DeMulHAC network demonstrates superior performance to other machine learning techniques.

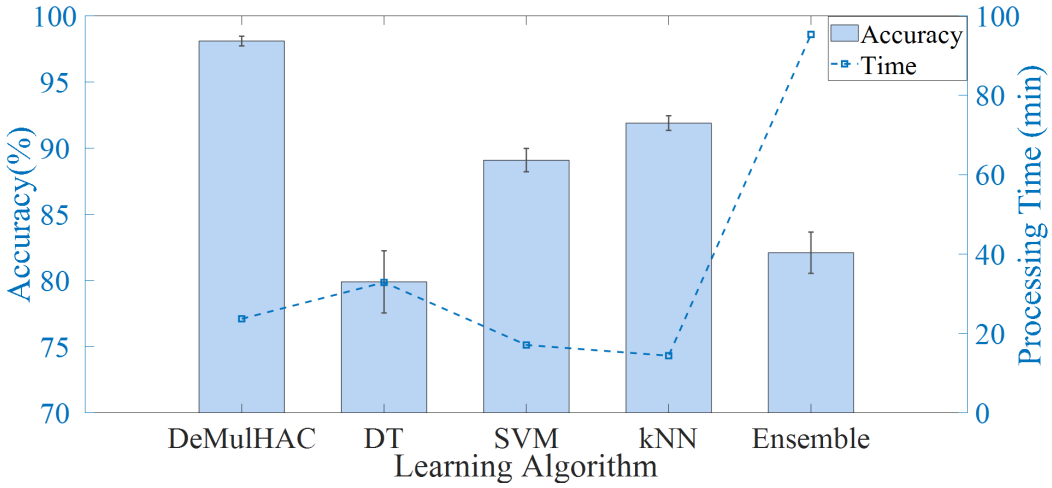


Figure 5. Comparison of DeMulHAC with ML techniques (on the ENABLE3S dataset).

The approach presented herein also implements automatic feature reduction using multi-factor ANOVA, which significantly reduces the computational complexity. It has been compared with another feature reduction technique, i.e., mRMR. Table 8 displays the results of DeMulHAC performance using both of these feature reduction methods. It shows that in comparison to mRMR, the set of features extracted from ANOVA performs better in terms of accuracy and MSE. Also, a significant difference exists between the training and testing accuracy of mRMR, which depicts that the network is overfitting [66]. This proves that the features extracted from ANOVA perform better on the proposed DeMulHAC.

Analyzing the classification results, we also interpret the significant positions for sensor placement by using the weights allotted to DeMulHAC's top 130 feature inputs. Results

(Table 9) show that 72 features are associated with IMUs, 51 with sEMG, and only 7 features with goniometers. The features listed in Table 9 are selected based on the weight allocation during the training process of the DeMulHAC. Specifically, p-values range from values significantly lower than 0.001 to 0.04, whereas the discarded ones are associated with p-values ranging from 0.06 to 0.90. It is worth highlighting, as the lower p-values are found to be related to the IMUs-based features. The features are extracted using Multi-factor ANOVA as expressed in (5):

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)ij + \epsilon ijk \quad (5)$$

where we have the following:

- Y_{ijk} is the observed value for the k-th observation in the i-th level of factor A and j-th level of factor B.
- μ is the overall mean of all observations.
- α_i is the effect of the i-th level of factor A.
- β_j is the effect of the j-th level of factor B.
- $(\alpha\beta)_{ij}$ is the interaction effect between factor A and factor B at the i-th and j-th levels.
- ϵ_{ijk} is the random error term associated with the k-th observation at the i-th level of A and j-th level of B.

These numerical data lead to the observation that IMUs are the most significant, sEMGs are second in importance, and goniometers are the least significant for activity classification.

Table 9 further provides insight into the body and muscle positions which can produce accurate classification results. It clearly displays that the position of the shank and thigh is more important than the waist IMU. Similarly, the tibialis anterior, biceps femoris, and semitendinosus muscles are more significant than the other muscles. The high relevance of the tibialis anterior (TA), biceps femoris (BF), and semitendinosus (ST) muscles, as highlighted in the feature selection analysis, aligns with their crucial roles in locomotion. For instance, the tibialis anterior is critical for dorsiflexion during the swing phase of gait, ensuring foot clearance [67]. Similarly, the biceps femoris and semitendinosus play vital roles in knee flexion and hip extension during the stance phase [68], contributing to stability and propulsion. These biomechanical roles make signals from these muscles particularly discriminative for identifying specific gait events and activities. Regarding sensor placements, the results underscore the significance of IMUs located on the shank and thigh, which are closely associated with the primary joints involved in movement, such as the knee and ankle. These placements capture dynamic variations in stride length, frequency, and joint angles, which are pivotal for accurate activity classification. In contrast, the relatively lower relevance of waist-mounted IMUs may stem from their reduced sensitivity to localized limb movements [26,69]. These insights not only enhance the interpretability of the model but also suggest practical guidelines for sensor placement in future studies. For example, focusing on lower limb sensors and muscle groups identified as significant in this study could improve model accuracy and generalizability when applied to other datasets or use cases, such as designing wearable systems for rehabilitation or sports performance monitoring.

Table 8. Comparison of mRMR with ANOVA.

Feature Reduction Technique	ANOVA	mRMR
Network Configuration	130, 65, 90	130, 65, 90
MSE	0.0081 ± 0.0011	0.01801 ± 0.0023
Training Accuracy (%)	99.3 ± 0.4590	95.57 ± 1.282
Validation Accuracy (%)	94.53 ± 0.4701	90.47 ± 1.115
Testing Accuracy (%)	95.4 ± 0.3994	86.29 ± 1.231
Overall Accuracy (%)	98.1 ± 0.3706	93.22 ± 1.147

Table 9. Feature ranking according to DeMulHAC weights.

Sensor Type	Body Part	Number of Features
IMU	Shank	35 features
	Thigh	26 features
	Waist	11 features
		72 features
sEMG	TA	13 features
	BF	10 features
	ST	9 features
	SOL	7 features
	RF	4 features
	MG	4 features
	VL	4 features
		51 features
Goniometer		7 features

This functional relevance suggests that these muscles are not only involved in basic locomotion but also in more complex activities requiring fine motor control and stability. The greater significance of these muscle groups compared to others could indicate their critical roles in distinguishing between different types of activities. For instance, the biceps femoris may show distinct activation patterns during running compared to walking, which could be leveraged in algorithm development for more precise activity recognition [70].

This specificity reinforces the notion that a targeted approach to sensor and muscle selection can enhance the classification accuracy of gait and activity recognition systems.

Additionally, the optimization of sensor arrangements, focusing on the shank and thigh while incorporating EMG data from key muscle groups, can lead to significant advancements in wearable technology. Such configurations can facilitate more accurate real-time monitoring of human movement [71], which is particularly beneficial in rehabilitation settings, where personalized feedback is essential for patient progress. The study confirms the hypothesis through a two-phase optimization strategy applied to the DNN. The results demonstrate that the DNN, specifically the Deep Multi-class Human Activity Classifier (DeMulHAC), achieves outstanding accuracy rates of 98.1% for activity classification and 99.96% for gait event detection. This is a notable improvement over prior methods: for instance, comparable models have achieved lower accuracy (such as around 92–94% with CNNs on simpler, single-sensor datasets). These performance metrics validate the initial hypothesis by showing that the optimized DNN architecture outperforms other models, such as SVMs and decision trees, particularly in handling the complexity and non-linearity of heterogeneous sensor data. Furthermore, the study highlights that the DNN model not only classifies with high accuracy but also achieves efficient processing times, which is crucial for real-time applications in wearable technology. The effective classification achieved by the DeMulHAC network underscores the importance of using a tailored architecture and a diverse sensor setup, as features from different sensor types, especially those from IMUs and selected muscles, are found to be critical for accurate classification. This optimization and the model's performance substantiate the hypothesis that a well-designed DNN with a multi-sensor input approach can meet the demands of precise human activity and gait event classification.

5.4. Limitation and Future Work

This study presents valuable insights into human activity classification but is subject to some imitations that need further exploration. First, the reliance on data from only ten healthy participants restricts the generalizability of the findings, as the dataset may not capture the variability present in more diverse populations or clinical scenarios. For instance, the inclusion of elderly individuals, those with mobility impairments, or participants from different demographic backgrounds could enhance the model's robustness

and applicability in real-world settings. Future studies should focus on acquiring datasets that reflect such diversity to improve the generalizability of the results. Second, the offline nature of the current classification approach limits its immediate applicability to real-time environments, such as rehabilitation or sports analytics. Transitioning the proposed model to real-time applications poses challenges, including reducing computational complexity and ensuring low-latency predictions. Potential strategies to address these challenges include model quantization, pruning, and optimization for deployment on edge devices. These approaches could enable real-time monitoring and analysis, making the system more practical for dynamic settings. Future research will prioritize these adaptations to expand the applicability of the model in clinical and everyday scenarios. Further investigations will integrate additional sensor modalities, test the model under varied sensor placements, and utilize advanced algorithms like RNNs, LSTMs, and ensemble methods to improve classification accuracy and robustness in diverse environments, including sports.

In conclusion, these findings emphasize the importance of the optimal placement of heterogeneous sensors and the integration of muscle-specific data. By refining these methodologies, we offer an innovative algorithm to improve the accuracy of activity classification systems, ultimately contributing to more effective interventions in both clinical and everyday settings.

6. Conclusions

This study introduces the DeMulHAC architecture, achieving 98.1 ± 0.37% accuracy for activity classification and 99.96 ± 0.05% for gait event detection. The ANOVA-based feature selection reduces features from 368 to 260 with an MSE of 0.0081 ± 0.0011, while highlighting the importance of IMU (72 of the top 130 features) and sEMG data from key muscles like the tibialis anterior and biceps femoris. Despite limitations, such as a small dataset and offline implementation, these results demonstrate the model's potential for real-time clinical and wearable applications. Future work will focus on expanding participant diversity and enabling real-time deployment. The proposed DeMulHAC architecture is validated through 10-fold cross-validation and comparisons with state-of-the-art methods, demonstrating superior accuracy and efficiency. ANOVA-based feature selection outperforms mRMR, with IMU and sEMG signals identified as the most critical for activity recognition.

Author Contributions: Conceptualization, T.J., A.R., S.Z., H.F.M., J.T., S.R. Data curation, T.J., A.R., J.T. and S.R. Formal analysis, T.J., A.R., S.Z.; Methodology, S.Z., A.R., H.F.M., J.T. and S.R. Software, J.T.; Supervision, A.R., J.T. and S.R. Writing—original draft, T.J., A.R., S.Z., H.F.M., J.T.; Writing—review & editing, T.J., A.R., S.Z., H.F.M., J.T. and S.R. All authors have read and agreed to the published version of the manuscript.

Funding: The presented work was supported by Higher Education Commission of Pakistan under NRPU-9540 project.

Data Availability Statement: ENABL3S dataset, used as input to this study, is already available at [8]; the code of the DNN-based classifier is available on the GitHub page of the corresponding author at https://github.com/dr-ali-raza/DeMulHAC (accessed on 25 November 2024).

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Maqbool, H.F.; Husman, M.A.B.; Awad, M.I.; Abouhossein, A.; Iqbal, N.; Dehghani-Sanij, A.A. A real-time gait event detection for lower limb prosthesis control and evaluation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *25*, 1500–1509. [CrossRef] [PubMed]
2. Deans, S.; Burns, D.; McGarry, A.; Murray, K.; Mutrie, N. Motivations and barriers to prosthesis users' participation in physical activity, exercise, and sport: A review of the literature. *Prosthetics Orthot. Int.* **2012**, *36*, 260–269. [CrossRef] [PubMed]
3. Jamieson, A.; Murray, L.; Stankovic, L.; Stankovic, V.; Buis, A. Human activity recognition of individuals with lower limb amputation in free-living conditions: A pilot study. *Sensors* **2021**, *21*, 8377. [CrossRef]
4. Bilal, M.; Rizwan, M.; Maqbool, H.F.; Ahsan, M.; Raza, A. Design optimization of powered ankle prosthesis to reduce peak power requirement. *Sci. Prog.* **2022**, *105*, 00368504221117895. [CrossRef]

5. Chen, Y.; Shen, C. Performance analysis of smartphone-sensor behavior for human activity recognition. *IEEE Access* **2017**, *5*, 3095–3110. [CrossRef]
6. Peng, F.; Peng, W.; Zhang, C.; Zhong, D. IoT assisted Kernel linear discriminant analysis based gait phase detection algorithm for walking with cognitive tasks. *IEEE Access* **2019**, *7*, 68240–68249. [CrossRef]
7. Zubair, M.; Ishaq, H.N.; Raza, A.; Maqbool, H.F.; Zafar, S. Development and evaluation of a low-cost data acquisition system using heterogeneous sensors. *Int. J. Sens. Netw.* **2022**, *40*, 45–56. [CrossRef]
8. Hu, B.; Rouse, E.; Hargrove, L. Benchmark datasets for bilateral lower-limb neuromechanical signals from wearable sensors during unassisted locomotion in able-bodied individuals. *Front. Robot. AI* **2018**, *5*, 14. [CrossRef]
9. Miller, J.D.; Beazer, M.S.; Hahn, M.E. Myoelectric walking mode classification for transtibial amputees. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 2745–2750. [CrossRef]
10. Goršič, M.; Kamnik, R.; Ambrožič, L.; Vitiello, N.; Lefeber, D.; Pasquini, G.; Munih, M. Online phase detection using wearable sensors for walking with a robotic prosthesis. *Sensors* **2014**, *14*, 2776–2794. [CrossRef]
11. Tucker, M.R.; Olivier, J.; Pagel, A.; Bleuler, H.; Bouri, M.; Lambercy, O.; Millán, J.D.; Riener, R.; Vallery, H.; Gassert, R. Control strategies for active lower extremity prosthetics and orthotics: A review. *J. NeuroEng. Rehabil.* **2015**, *12*, 1–30. [CrossRef] [PubMed]
12. Lara, O.D.; Labrador, M.A. A survey on human activity recognition using wearable sensors. *IEEE Commun. Surv. Tutorials* **2012**, *15*, 1192–1209. [CrossRef]
13. Chen, B.; Zheng, E.; Wang, Q. A locomotion intent prediction system based on multi-sensor fusion. *Sensors* **2014**, *14*, 12349–12369. [CrossRef]
14. Preece, S.J.; Goulermas, J.Y.; Kenney, L.P.; Howard, D.; Meijer, K.; Crompton, R. Activity identification using body-mounted sensors—A review of classification techniques. *Physiol. Meas.* **2009**, *30*, R1–R33. [CrossRef]
15. Ramasamy Ramamurthy, S.; Roy, N. Recent trends in machine learning for human activity recognition—A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1254. [CrossRef]
16. Hussain, T.; Maqbool, H.F.; Iqbal, N.; Khan, M.; Salman; Dehghani-Sanij, A.A. Computational model for the recognition of lower limb movement using wearable gyroscope sensor. *Int. J. Sens. Netw.* **2019**, *30*, 35–45. [CrossRef]
17. Hussain, T.; Iqbal, N.; Maqbool, H.F.; Khan, M.; Awad, M.I.; Dehghani-Sanij, A.A. Intent based recognition of walking and ramp activities for amputee using sEMG based lower limb prostheses. *Biocybern. Biomed. Eng.* **2020**, *40*, 1110–1123. [CrossRef]
18. Culhane, K.M.; Lyons, G.M.; Hilton, D.; Grace, P.A.; Lyons, D. Long-term mobility monitoring of older adults using accelerometers in a clinical environment. *Clin. Rehabil.* **2004**, *18*, 335–343. [CrossRef]
19. Mannini, A.; Sabatini, A.M. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* **2010**, *10*, 1154–1175. [CrossRef]
20. Najafi, B.; Aminian, K.; Paraschiv-Ionescu, A.; Loew, F.; Bula, C.J.; Robert, P. Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly. *IEEE Trans. Biomed. Eng.* **2003**, *50*, 711–723. [CrossRef]
21. Demrozi, F.; Pravadelli, G.; Bihorac, A.; Rashidi, P. Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey. *IEEE Access* **2020**, *8*, 210816–210836. [CrossRef] [PubMed]
22. Karpiński, R.; Krakowski, P.; Jonak, J.; Machrowska, A.; Maciejewski, M. Comparison of Selected Classification Methods Based on Machine Learning as a Diagnostic Tool for Knee Joint Cartilage Damage Based on Generated Vibroacoustic Processes. *Appl. Comput. Sci.* **2023**, *29*, 136–150. [CrossRef]
23. Thakur, D.; Dangi, S.; Lalwani, P. A novel hybrid deep learning approach with GWO—WOA optimization technique for human activity recognition. *Biomed. Signal Process. Control* **2025**, *99*, 106870. [CrossRef]
24. Brzenczek, C.; Klopfenstein, Q.; Hähnel, T.; Fröhlich, H.; Glaab, E. Integrating digital gait data with metabolomics and clinical data to predict outcomes in Parkinson's disease. *npj Digit. Med.* **2024**, *7*, 235. [CrossRef]
25. Rafique, W.; Naqvi, S.M.; Chambers, J.A. Mixed source prior for the fast independent vector analysis algorithm. In Proceedings of the 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), Rio de Janeiro, Brazil, 10–13 July 2016; pp. 1–5.
26. Attal, F.; Mohammed, S.; Dedabrishvili, M.; Chamroukhi, F.; Oukhellou, L.; Amirat, Y. Physical human activity recognition using wearable sensors. *Sensors* **2015**, *15*, 31314–31338. [CrossRef] [PubMed]
27. Khairy, D.; Behery, G.; Ewees, A.A.; AbdElrazek, E. Performance Evaluation of Neural Network and Deep Neural Network for Human Activity Recognition. *Int. J. Comput. Appl.* **2018**, *180*, 44–50.
28. Hayashi, Y. Does deep learning work well for categorical datasets with mainly nominal attributes? *Electronics* **2020**, *9*, 1966. [CrossRef]
29. Gumaei, A.; Hassan, M.M.; Alelaiwi, A.; Alsalman, H. A hybrid deep learning model for human activity recognition using multimodal body sensing data. *IEEE Access* **2019**, *7*, 99152–99160. [CrossRef]
30. Yan, W.; Li, G.; Wu, Z.; Wang, S.; Yu, P.S. Extracting diverse-shapelets for early classification on time series. *World Wide Web* **2020**, *23*, 3055–3081. [CrossRef]
31. Wu, Z.; Shen, S.; Lian, X.; Su, X.; Chen, E. A dummy-based user privacy protection approach for text information retrieval. *Knowl.-Based Syst.* **2020**, *195*, 105679. [CrossRef]
32. Wu, Z.; Zhu, H.; Li, G.; Cui, Z.; Huang, H.; Li, J.; Chen, E.; Xu, G. An efficient Wikipedia semantic matching approach to text document classification. *Inf. Sci.* **2017**, *393*, 15–28. [CrossRef]

33. Helmi, A.M.; Al-Qaness, M.A.; Dahou, A.; Damaševičius, R.; Krilavičius, T.; Elaziz, M.A. A novel hybrid gradient-based optimizer and grey wolf optimizer feature selection method for human activity recognition using smartphone sensors. *Entropy* **2021**, *23*, 1065. [CrossRef]
34. Kiran, S.; Khan, M.A.; Javed, M.Y.; Alhaisoni, M.; Tariq, U.; Nam, Y.; Damaševičius, R.; Sharif, M. Multi-Layered Deep Learning Features Fusion for Human Action Recognition. *Comput. Mater. Contin.* **2021**, *69*, 4061–4075. [CrossRef]
35. Huang, W.; Zhang, L.; Wu, H.; Min, F.; Song, A. Channel-Equalization-HAR: A light-weight convolutional neural network for wearable sensor based human activity recognition. *IEEE Trans. Mob. Comput.* **2022**, *22*, 5064–5077. [CrossRef]
36. Tang, Y.; Zhang, L.; Min, F.; He, J. Multiscale deep feature learning for human activity recognition using wearable sensors. *IEEE Trans. Ind. Electron.* **2022**, *70*, 2106–2116. [CrossRef]
37. Moencks, M.; De Silva, V.; Roche, J.; Kondoz, A. Adaptive feature processing for robust human activity recognition on a novel multi-modal dataset. *arXiv* **2019**, arXiv:1901.02858.
38. Cheng, X.; Zhang, L.; Tang, Y.; Liu, Y.; Wu, H.; He, J. Real-time human activity recognition using conditionally parametrized convolutions on mobile and wearable devices. *IEEE Sensors J.* **2022**, *22*, 5889–5901. [CrossRef]
39. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
40. Calatrava-Nicolás, F.M.; Mozos, O.M. Deep adversarial learning with activity-based user discrimination task for human activity recognition. *arXiv* **2024**, arXiv:2410.12819.
41. Sagha, H.; Digumarti, S.T.; del R. Millán, J.; Chavarriaga, R.; Calatroni, A.; Roggen, D.; Tröster, G. Benchmarking classification techniques using the Opportunity human activity dataset. In Proceedings of the 2011 IEEE International Conference on Systems, Man, and Cybernetics, Anchorage, AK, USA, 9–12 October 2011; pp. 36–40. [CrossRef]
42. El-Gayar, O.F.; Hossain, M.I. Metaheuristics: From design to implementation. *IEEE Access* **2020**, *8*, 127142–127167. [CrossRef]
43. Rebelo, D.; Amma, C.; Gamboa, H.; Schultz, T. Human activity recognition for an intelligent knee orthosis. In Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, Barcelona, Spain, 11–14 February 2013; Volume 2, pp. 368–371.
44. Hammerla, N.Y.; Halloran, S.; Plötz, T. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv* **2016**, arXiv:1604.08880.
45. San, P.P.; Kakar, P.; Li, X.L.; Krishnaswamy, S.; Yang, J.B.; Nguyen, M.N. Deep learning for human activity recognition. In *Big Data Analytics for Sensor-Network Collected Intelligence*; Morgan Kaufmann: Burlington, MA, USA, 2017; pp. 186–204.
46. Xu, X.; Li, W.; Ran, Q.; Du, Q.; Gao, L.; Zhang, B. Multisource remote sensing data classification based on convolutional neural network. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 937–949. [CrossRef]
47. Zhang, M.; Li, W.; Du, Q.; Gao, L.; Zhang, B. Feature extraction for classification of hyperspectral and LiDAR data using patch-to-patch CNN. *IEEE Trans. Cybern.* **2018**, *50*, 100–111. [CrossRef] [PubMed]
48. Ignatov, A. Real-time human activity recognition from accelerometer data using convolutional neural networks. *Appl. Soft Comput.* **2018**, *62*, 915–922. [CrossRef]
49. Moya Rueda, F.; Grzeszick, R.; Fink, G.A.; Feldhorst, S.; Ten Hompel, M. Convolutional neural networks for human activity recognition using body-worn sensors. *Informatics* **2018**, *5*, 26. [CrossRef]
50. Hassan, M.M.; Huda, S.; Uddin, M.Z.; Almogren, A.; Alrubaian, M. Human activity recognition from body sensor data using deep learning. *J. Med Syst.* **2018**, *42*, 1–8. [CrossRef]
51. Zhe, S.; Naqvi, S.A.; Yang, Y.; Qi, Y. Joint network and node selection for pathway-based genomic data analysis. *Bioinformatics* **2013**, *29*, 1987–1996. [CrossRef]
52. Meyer-Bäse, A. *Pattern Recognition and Signal Analysis in Medical Imaging*; Academic Press: Cambridge, MA, USA, 2004.
53. Bejani, M.; Gharavian, D.; Charkari, N.M. Audiovisual emotion recognition using ANOVA feature selection method and multi-classifier neural networks. *Neural Comput. Appl.* **2014**, *24*, 399–412. [CrossRef]
54. Gao, W.; Hu, L.; Zhang, P. Class-specific mutual information variation for feature selection. *Pattern Recognit.* **2018**, *79*, 328–339. [CrossRef]
55. Liu, G.; Zhang, L.; Han, B.; Zhang, T.; Wang, Z.; Wei, P. sEMG-based continuous estimation of knee joint angle using deep learning with convolutional neural network. In Proceedings of the 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), Vancouver, BC, Canada, 22–26 August 2019; pp. 140–145.
56. De-La-Hoz-Franco, E.; Ariza-Colpas, P.; Quero, J.M.; Espinilla, M. Sensor-based datasets for human activity recognition—A systematic review of literature. *IEEE Access* **2018**, *6*, 59192–59210. [CrossRef]
57. Huang, G.B. Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Trans. Neural Netw.* **2003**, *14*, 274–281. [CrossRef] [PubMed]
58. Charilaou, P.; Battat, R. Machine learning models and over-fitting considerations. *World J. Gastroenterol.* **2022**, *28*, 605–607. [CrossRef] [PubMed] [PubMed Central]
59. Cheng, Y.; Wang, D.; Zhou, P.; Zhang, T. Model compression and acceleration for deep neural networks: The principles, progress, and challenges. *IEEE Signal Process. Mag.* **2018**, *35*, 126–136. [CrossRef]
60. Møller, F.M. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **1993**, *6*, 525–533. [CrossRef]
61. Twomey, J.M.; Smith, A.E. Performance measures, consistency, and power for artificial neural network models. *Math. Comput. Model.* **1995**, *21*, 243–258. [CrossRef]

62. Hunter, D.; Yu, H.; Pukish, M.S., III; Kolbusz, J.; Wilamowski, B.M. Selection of Proper Neural Network Sizes and Architectures—A Comparative Study. *IEEE Trans. Ind. Informatics* **2012**, *8*, 228–240. [CrossRef]
63. Kılıçarslan, S.; Adem, K.; Çelik, M. An overview of the activation functions used in deep learning algorithms. *J. New Results Sci.* **2021**, *10*, 75–88. [CrossRef]
64. Babani, L.; Jadhav, S.; Chaudhari, B. Scaled Conjugate Gradient Based Adaptive ANN Control for SVM-DTC Induction Motor Drive. In *Artificial Intelligence Applications and Innovations*; Iliadis, L., Maglogiannis, I., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 384–395. [CrossRef]
65. Taborri, J.; Palermo, E.; Rossi, S. WARNING: A Wearable Inertial-Based Sensor Integrated with a Support Vector Machine Algorithm for the Identification of Faults during Race Walking. *Sensors* **2023**, *23*, 5245. [CrossRef]
66. Zhu, Y.; Xiong, Y. Towards data science. *Data Sci. J.* **2015**, *14*, 1–7. [CrossRef]
67. Maharaj, J.N.; Cresswell, A.G.; Lichtwark, G.A. Tibialis anterior tendinous tissue plays a key role in energy absorption during human walking. *J. Exp. Biol.* **2019**, *222*, jeb191247. [CrossRef]
68. Hirose, N.; Tsuruike, M.; Higashihara, A. Biceps femoris muscle is activated by performing nordic hamstring exercise at a shallow knee flexion angle. *J. Sport. Sci. Med.* **2021**, *20*, 275. [CrossRef] [PubMed]
69. Taborri, J.; Palermo, E.; Rossi, S. Automatic detection of faults in race walking: A comparative analysis of machine-learning algorithms fed with inertial sensor data. *Sensors* **2019**, *19*, 1461. [CrossRef] [PubMed]
70. Sasaki, K.; Neptune, R.R. Differences in muscle function during walking and running at the same speed. *J. Biomech.* **2006**, *39*, 2005–2013. [CrossRef] [PubMed]
71. Palumbo, F.; Gallicchio, C.; Pucci, R.; Micheli, A. Human activity recognition using multisensor data fusion based on reservoir 746 computing. *J. Ambient. Intell. Smart Environ.* **2016**, *8*, 87–107. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.