# EEG-Driven Physical Load Detection During Human Walking for Robotic and Automation Systems

*Harsha* Durairaj[1], *M.* Jagannath[2], U*ma* Parvathi G[1], *Saravanakumar* Jaganathan[3,4,5], *K.* Adalarasu[6*], *K.* Narasimhan[6], *Rahul* Bhardwaj[7]

[1]Department of Robotics & Artificial Intelligence, School of Electrical and Electronics Engineering, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India
[2]School of Electronics Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India
[3]Institute of Research and Development, Duy Tan University, Da Nang, Vietnam
[4]School of Engineering & Technology, Duy Tan University, Da Nang, Vietnam
[5]Biomedical Engineering Research Group, School of Engineering, University of Leicester, Leicester LE1 7RH, United Kingdom
[6]School of Electrical and Electronics Engineering, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India
[7]Chief Technical Officer, Qneuro India Private Limited, Chennai, Tamil Nadu, India

**Abstract.** Understanding neural responses to varying physical loads is essential for developing ergonomic designs. Conventional methods for analyzing peripheral muscle activity, such as electromyography (EMG) and kinematic analysis, provide only limited insight into the cortical dynamics associated with physical tasks. To overcome this limitation, the present study introduces an electroencephalography (EEG)-based approach to investigate brain activity during load-bearing conditions. Participants performed a 100-meter walking task while carrying a 5 kg shoulder load, during which raw EEG signals were recorded. These signals were transformed using Continuous Wavelet Transform (CWT) to generate scalograms, capturing both temporal and frequency-domain characteristics of neural activity. Deep learning (DL) models were then trained, validated, and tested using these representations, and their performance was evaluated through standard metrics. Several DL architectures, including CNN, ResNet18, VGG19, DenseNet, and ResNet50, were employed to extract spatial–temporal features associated with load conditions. Among these, ResNet18 achieved the highest accuracy of 66.83%, outperforming conventional feature-based approaches. Additionally, the occipital cortex showed the highest classification accuracy (69.09%) in distinguishing between no-load and 5 kg load conditions. These findings highlight the potential of DL-based EEG analysis for workload monitoring, fatigue assessment, and brain–computer interface applications.

*****Corresponding author: adalarasu@eie.sastra.edu

## 1 Introduction

Knowledge about the human brain's reaction to changes in physical workload is needed for advancements in occupational safety, ergonomic system development, and rehabilitation approaches. Exercise also affects muscles and biomechanics. This involves complex neural mechanisms for organising motor planning, movement, coordination, and perceiving fatigue. Traditional techniques like electromyography (EMG) and motion analysis mainly measure peripheral responses and provide minimal insight into the cortical mechanisms of exertion. Thorough knowledge of the neural processes that regulate intrinsic physical load is essential for studies in human-machine interaction. Moderate to intense physical exertion alters coordination between the peripheral and central nervous systems, affecting motor planning, coordination, and fatigue control. Taori and Lim [13] suggested conventional evaluative methodologies, such as EMG; Refai et al. [9] developed force-sensing and musculoskeletal modelling methods; Zhou et al. [20] characterised wearable tactile sensing systems. These systems primarily record muscular or biomechanical outputs and therefore provide limited access to the neural correlates of exertion. Conversely, electroencephalography (EEG) is a non-invasive method for measuring cortical activity, providing insight into cerebral dynamics during movement and effort modulation. Empirical support exists for this utility: Ramadan et al. [10] found significant EEG activity changes related to physical fatigue during manual tasks. Brouwer et al. [2] identified differing cortical patterns depending on workload, using EEG spectral power and event-related potentials. Xiao et al. [17] demonstrated that EEG power responses are sensitive to lifting-induced fatigue. Together, these findings confirm EEG as an effective tool for capturing neural reactions to various physical loads. The current research uses EEG, time-frequency analysis, and deep learning to decode cortical activity in response to different physical loading states during walking.

## 2 Literature Review

With the recent implementation of deep learning, in particular, convolutional neural networks (CNNs), the EEG analysis has been enhanced from manual extraction of features to the automatic extraction of a complex spatiotemporal pattern. A review of deep learning approaches to EEG analysis by Craik et al. [3] emphasized that end-to-end CNN-based systems are effective in the learning of dynamic neural representations of fatigue and workload detection. Tabar et al. [14] established that converting one-dimensional EEG time-series signals into structured representations apply CNN architectures to train various features of neural data successfully. Mohamed et al. [10] proved that machine learning and time-frequency representations to detect fatigue through the study of physiological signals have yielded better classification results when spectral data is included.

Deep Learning-based EEG fatigue detection, as was introduced by Ren et al. [12], has also been successfully applied in other areas of application and proved to be robust in the face of different operational circumstances. Recent works in EEG-based deep learning systems with increasingly advanced attention mechanisms and architectures, which are based on a graph. The reasoning behind understanding that structured neural representations can be learned effectively using multi-channel EEG data was presented by Dai et al. [4], who suggested a deep learning framework for EEG motor imagery classification. The review of deep learning methods to study EEG by Craik et al. [3] improved the significance of the data-efficient representation learning concept to enhance robustness and generalisation, especially in the context of using limited training data, which is more pertinent to the EEG-based load classification.

Beyond EEG, numerous studies have been carried out in load detection and classification based on physiological sensor modalities or wearable sensors. As indicated by Mohamed et al. [10], the research on physiological signal-based fatigue detection by machine learning proved that neural network models can classify fatigue levels by analysing physiological and biomechanical data, which makes them possible to be used in the process of workload evaluation. Low-back electromyography (EMG) data were used by Totah et al. [15] to categorise dynamic lifting loads as they provide a biomechanical basis that integrates muscle activity patterns and exertion levels. This study was further expanded by Gorsic et al. [5], who utilized wearable IMU and EMG devices to identify the position of load and weight during gait, emphasising that various approaches could be applicable in a real-world configuration to recognise loads. In accordance with the methodologies, Tabar et al. [14] determined the use of EEG measures of lifting activity and transformed them into continuous wavelet transform (CWT) scalograms to effectively record the transient and frequency-modulated neural responses, which follows the recent developments.

EEG classification using CNN, in particular, when combined with scalograms obtained using CWT, enhances the responsiveness and interpretability of assistive technologies (including robotic exoskeletons, artificial limbs, and neurorehabilitation systems). The CNN framework, aimed at relating neural activity to measurable exertion, was described by Brouwer et al. [2], Ahn et al. [3], and Taori and Lim [13], which is in line with the principles. The above-discussed methods provide an assurance that they could be applied in various fields and include finding loads, measuring fatigue, assessing ergonomics, managing rehabilitation, enhancing sports performance, and muscle activation studies. Electromyographic (EMG) data combined with biomechanical analyses have been shown by recent works by Boschman et al. [1] and Totah et al. [15] published in Applied Ergonomics to be beneficial in providing extensive workload assessment using the electroencephalogram (EEG)-based analytical framework. Also, combining different physiological measurements, such as EMG signals, heart rate measurements, and inertial, has been proposed to increase the resilience and adaptability of such systems as demonstrated in the study of Taori and Lim [13], Refai et al. [9], and Zhou et al. [20]. EEG-based deep learning frameworks not only monitor physical workloads but also provide considerable opportunities for the development of Brain Computer Interface (BCI) technologies.

The studies by Ahn et al. [3], Brouwer et al. [2], and Zhang et al. [19] discussed that BCIs may progress to more sophisticated interfaces that are able to adapt efficiently to the physical and cognitive state of the participant. Moreover, Aharon et al. [7] had a model under which the principles of learning are useful in deriving the high-dimensional data in optimum and discriminative features, which provides a theoretical underpinning to deep learning-based EEG and BCI systems. In this way, further studying of cortical reactions to new physical loads preconditions building of this groundwork in the future BCIs, which will be able to anaesthetically predict the intentions of users, and provide them with a smart and situation-dependent assistance in real time. Thus, the presented analysis can improve not only the classification of physical loads with the help of EEG but also play an important role in the broader goal of creating adaptive, human-centric, and neurophysiology-inspired-human-machine interaction systems. The deep learning structures based on EEG have significant ramifications in the field of Brain-Computer Interface (BCI) systems, beyond the field of physical load monitoring. Ahn et al. [3], Brouwer et al. [2], and Zhang et al. [19] showed that their BCIs may be advanced to the adaptive systems able to react to physical and cognitive conditions of the users by interpreting the cortical activity linked to the effort and fatigue during the motor tasks. While prior studies have primarily examined the EEG signal for cognitive workload or fatigue assessment, little research has focused on physical load detection using EEG. Most existing studies to predict/classify the backpack's

different load carrying using EMG signal during real-time gait movement, leverage the neural cortical mechanisms underlying physical exertion largely unexplored. To the best of the authors' knowledge, this research is among the first to investigate the classification of various shoulder load carrying while walking on a plan floor using EEG biomarkers through Deep learning techniques. This contribution unifies the gap between modern AI methodologies and robotic load-carrying motion analysis using brain signals while walking, offering meaningful advancements in this interdisciplinary field.

A novel approach is proposed that utilises real-time EEG data to accurately classify various shoulder loads during gait activity, as shown in Figure 1.

The key contributions of this study include

- First-time analysis of neural signal to classify different types of load-carrying weight while real-time gait
- Designed own and transfer learning model, evaluated models (CNN, ResNet18, VGG19, DenseNet and ResNet50) for classification of subject back bag load weight.
- Modelling of key parts of the brain that distinguish the condition of shoulder load.

The rest of this paper is organised in the following way: Section 2 provides a description of the materials and methods that include research design, participant selection, experimental design, EEG data collection, and feature extraction procedure. Section 3 shows the outputs of deep learning structures, which are implemented to classify different shoulder load-carrying actions. In Section 4, the findings are discussed within the framework of the existing neuroscience literature with a focus on the neural mechanisms that accompany various load-carrying activities. Finally, Section 5 wraps up the paper to summarise the main results, their implications for robotic load-carrying systems and suggest future research directions by focusing on the real-time directions as opportunities for real-time EEG-based BCI.

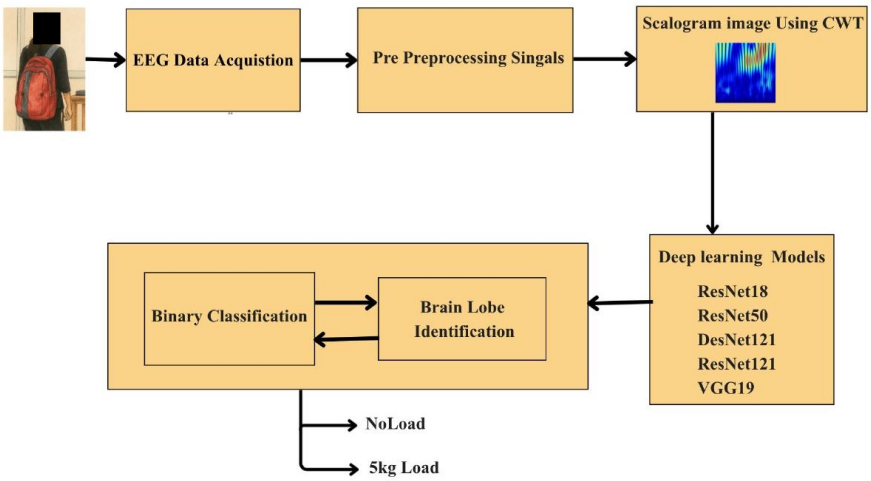



**Fig. 1.** Methodologies used to classify shoulder load carrying during walking

## 3 Materials and Methods

This session provides an overview of the particulars of the matter under study, the experimental framework, the experimental setup to gain data, the data processing, and the techniques of data gathering to classify the physical load levels in terms of the EEG measurements. The research objective of the experiments was to test the neural correlates

that were associated with physical effort during controlled activities that involved back load carriage during ambulatory conditions in the absence of load and a constant load of 5 kg, and used real-time EEG data recording.

### 3.1 Subjects Summary

This study was ethically approved by the Ethical Review Board of Our University, with participation being purely voluntary. Although the participants personally gave verbal informed consent before the commencement of the study, guardians of all the participants signed the informed consent letter after being made fully aware of what the study involved. The study was conducted in accordance with the ethical requirements put in place by the university. A systematic method was adopted to ensure the accuracy in collecting and evaluating data. The study involved twelve healthy residents (six males and six females, aged 19 to 21 years) who volunteered to participate in the research. The mean height and weight were 163.29 cm (SD = 6.52) and 53.41 kg (SD = 7.14), respectively. Every respondent was a right-handed individual with no neurological or musculoskeletal problems. Participants had been thoroughly briefed on the purpose of the study and methodologies before the session started.

### 3.2 Cognitive Health Assessment

Regarding the assessment of cognitive health, all the participants completed the Mini-Mental State Examination (MMSE) before the experiment. Different cognitive areas, including language, orientation, memory, responsiveness and visuospatial, were assessed in this examination with a rating scale of 0 - 30. To guarantee normal cognitive functioning, a threshold score of more than 25 was fixed to ensure that the participants were eligible for EEG data collection.

### 3.3 Load Stimulus

Two load conditions were studied: the no-load condition and the load condition that included an extra weight of 5 kg. In the no-load condition, the participants performed a walking activity without a load. In the load case, in their turn, they carried a typical dual-strap backpack with books, and this weighed 5 kg. This force can be compared to the weight of about 10-12 per cent of the mean body weight of participants used, and this is a standard commonly applied in ergonomic research to model the levels of intermediate physical workload. Within both experimental conditions, each of the no-load and 5 kg load participants was asked to walk barefoot over a flat and straight path that was 10 meters long, and the participants were asked to do this five consecutive times; each condition took about 3 minutes.

Each 3-minute gait with load carrying session was spaced with a 3-minute rest period, with randomized order to mitigate order effects. Throughout the experiment, EEG data were continuously recorded using scalp-mounted equipment to analyse neural responses. The experimental protocol is summarized in Figure 2.

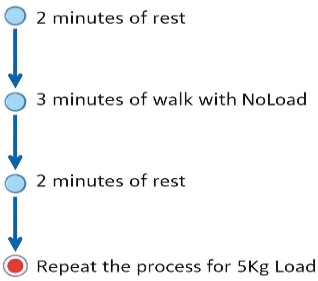



**Fig. 2.** Overview of the experimental protocol, including load-carrying sessions, rest, and EEG data recording

### 3.4 EEG Data Acquisition and Preprocessing

Hippocampus plays a role in the memory consolidation process, learning, and spatial processing, sensory perception, navigation, and sensory perception. Although the fact that transient memory functions are largely associated. The mechanism of long-term memory in the prefrontal cortex involves the hippocampus, neocortex, and amygdala. Electroencephalogram (EEG) data were used in the current study, recorded with an 8-channel Neuro-State Hat by Qneuro, Inc., a wireless multichannel instrument that is known to have a superior resolution in time and accuracy.


**Fig. 3.** QNeuro Neuro-State Hat : (A). Side view of the Hat, (B) Electrodes of the Hat, (C) Reference electrode of the Hat placed on the earlobe

The electrodes were to be placed as per the 10-20 international electrode placement. Figure 3, a system that enables the scalp to be covered extensively. Furthermore, reference electrodes were placed at the ears in order to enhance the quality of the recorded signals. The Neuro-State Hat has a time resolution and electrode arrangement, filtered to meet the motion of the electrical activity of the brain. Eight bolts were fitted in the cap: one reference electrode and active electrodes, which are aligned according to the 10-20 electrode system. This system permitted real-time evaluation of contact quality and electrode impedance, as reflected in colour-coded feedback in the data acquisition program. The Q-neuro application interface displays coloured indicators to assess the quality of the electrode contact.

- Green: Indicates optimal contact quality with impedance levels that are adequate for precise recordings (greater than 90% signal quality).
- Red: Represents poor contact quality, indicating that an adjustment is necessary. –
- Orange: Signifies average contact quality, implying that better positioning is required (as illustrated in Figure 4).


**Fig. 4.** Real-time signal strength displayed on QNeuro Analytics

Data collection in the electroencephalography (EEG) was conducted in a soundproof room. laboratory to minimise outside influences and guarantee the accuracy of the information. EEG recordings were acquired using the QNeuro Analytics System that incorporates an 8-channel, active EEG cap. Eight electrodes were used, including Fp1, Fp2, O1, O2, FC3, Cz, Fz, FC4 on the head and detailed electrode specifications are provided in Table 1. The left ear lobes were also subjected to scalp according to the 10/20 system, with one sensor reference in the controlled environment and the test subjects in the gait with an EEG headset. Figure 4 shows the real-time signal strength displayed on QNeuro Analytics. As illustrated in Figure 6, the experimental setup enabled real-time brain activity capture throughout the load-carrying task.

All electrodes were checked to exhibit constant (green) contact before every recording session. status. In case of poor connections with any electrode channel, then adjustments were reapplied to the gel or made until satisfactory impedance values were obtained. Ongoing censorship helped to ensure the high stability of signal quality throughout the session. EEG recording was done at a rate of 250 Hz/channel and entered into the local database of the system to be analysed later from the raw time series of EEG. The data were then exported to undergo processing of external signals and feature extraction. Custom analysis programs were written in Python.

To reduce external influences and guarantee data integrity, the EEG data collection experiment was conducted with no acoustic disturbances. Before data collection, all participants completed a demographic information sheet that recorded details such as name, age, education, and gender. On a 10-meter track, each participant made five continuous runs with no load and a 5 kg load condition (randomized order), with 3 minutes rest time provided as shown in Figure 5.

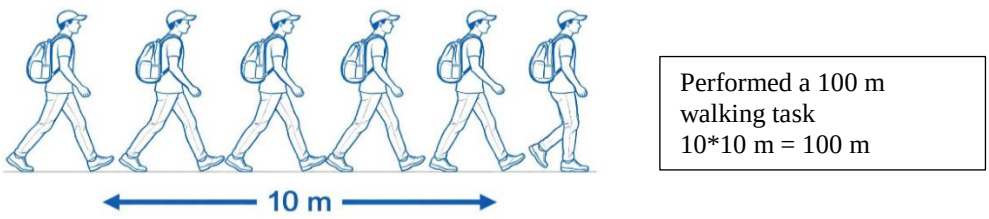
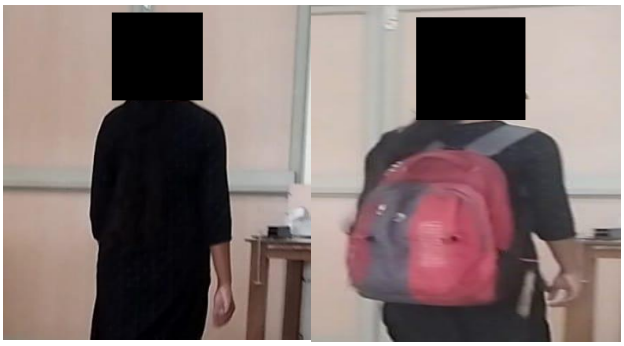



**Fig. 5**. Representation of human load (no load and 5 kg) carrying activity walking paths.

**Table 1.** Specifications of electrode details

| Electrode | Specification |
|---|---|
| EEG Channels | 8 EEG channels + 1 Reference channel + 1 Bias Channel |
| Reference Electrode Position | Ear |
| Channel Electrode Position | Fp1, Fp2, O1, O2, FC3, Cz, Fz, FC4 |
| EEG Electrode Materials | Hybrid of non-contact EEG electrode and Dry electrode plated with gold |
| Weight | 260g |



A B

**Fig. 6.** Experimental setup and EEG data acquisition with the Neuro-State Hat device during gait with different load carrying conditions: (A). No Load (B). 5 kg Load

### 3.5 EEG Signal Preprocessing

The raw EEG signals that were recorded by the QNEURO system were previously subject to the extraction of features pre-processing in order to enhance the signal's quality and remove artifacts. This preparation was done. written with Python, and utilizing libraries like NumPy, Pandas, SciPy, Matplotlib, etc. and PyWavelets. The steps that were involved in the entire pipeline included filtering, segmentation, time frequency transformation, artifact rejection, and normalization. The raw EEG waveforms represent time variations in the statistical properties of the waveforms and are often subject to random noise. To enhance the quality of the signals and guarantee data reliability, the EEG data was pre-processed in large amounts.

This stage used both notch and band-pass filters to filter the noise but preserve the necessary signal components. A 4th-order Butterworth band-pass filter with a cutoff frequency range of the EEG that was desired was kept to 0.4-40 Hz. Furthermore, a notch powerline interference was removed by using a filter set at 49-51 Hz. The initial and final 10 seconds of every recording were cut to eliminate any non-persistent effects at the beginning and the end of the tasks. The window function is used in the proposed algorithm as a mathematical tool that sets to zero values with an interval being selected, which splits

the original data into multiple consecutive sets. In particular, rectangular window functionality was used to separate the original. Signal was added to non-overlapping 5-second periods (150 seconds), causing 130 seconds one segment after another. The use of the time-windowing technique helped in exploiting smaller signal segments, which increases prediction accuracy.

Each epoch was normalized channel-wise to a range of [-1, 1] using min–max normalization to ensure uniform amplitude scaling across all electrodes, and this ensures that features contribute equally to machine learning models. For automatically removing artifacts, previous methods used channel-based statistical thresholding approaches to eliminate abnormal activities, as reported in EEG fatigue and workload studies by Delorme et al. [18], or adaptive filtering techniques employing reference channels to regress targeted artifacts, as discussed in EEG workload analysis by Brouwer et al. [2]. Independent component analysis (ICA) is a well-known approach for separating artifacts from neural signals, as reviewed in EEG-based brain–computer interface literature by Lotte et al. [6]. However, ICA-based methods have limited effectiveness in removing transient, non-biological artifacts such as impedance changes caused by headset motion and are computationally expensive, making them more suitable for offline analysis. In the present work, ASR techniques were employed for artifact correction using the QNeuro EEG cap system, which incorporates an in-built ASR algorithm for removing artifacts from recorded EEG signals. After preprocessing and normalization, Continuous Wavelet Transform (CWT) was applied using the Morlet wavelet with frequency range 0.4 – 45 Hz and across 128 scales to generate time–frequency scalogram images for each EEG epoch. The resulting scalograms (size ≈256 × 256 pixels) visually represent the temporal and spectral dynamics of EEG activity across all eight electrodes (F1, F2, O1, O2, FC3, FC4, Fz, and Cz) and were stored in respective output folders for use as inputs to a deep learning CNN-based classification model.

### 3.6 Deep Learning Models

In this study, EEG scalogram-based classification was employed to differentiate between no-load and 5 kg load conditions. For limited observational data, careful selection and optimization of convolutional neural network (CNN) architectures were critical to ensure robust feature extraction and prevent overfitting. In the present study, we use DenseNet121, ResNet18, ResNet50, VGG19, and a custom fine-tuned CNN for comparison.

#### *3.6.1DenseNet-121*

DenseNet-121 has 4 dense blocks arranged as [19, 6], meaning 120 convolutional layers in total. Each convolution is preceded by Batch Normalization and ReLU, giving 120 BN + 120 ReLU layers. Between dense blocks are 3 transition layers, each including 1×1 convolution + Average Pooling for down-sampling. At the end, there is a final Global Average-Pooling layer and 1 Fully Connected output layer.

DenseNet121 utilizes dense connectivity, in which each layer receives feature maps from all preceding layers. The DenseNet layer equation is represented in Eq. (1):

$$x_l = H_l\ ((x_0, x_1, \dots\dots\dots x_l)) \qquad (1)$$

Here, $H_l$ represents the operations performed at the layer $l$ (batch normalization → ReLU → convolution), and [·] denotes the concatenation of feature maps. This design encourages:

- Feature reuse, allowing deeper layers to access low-level features directly.
- Mitigation of the vanishing gradient problem by enabling shorter paths for gradient flow.

For binary classification, the original DenseNet classifier was replaced with a custom head. The model employs a two-layer ReLU-based architecture (Eq. (2) similar to that analyzed by Mahmud et al. [8]:

$$\hat{y}_i = \sigma(\omega_2 \,\mathrm{Re\,LU}(\omega_1 f_i + b_1) + b_2) \qquad (2)$$

where $f_i$ is the flattened feature vector, $\omega_1, \omega_2$ and $b_1$, $b_2$ are trainable weights and biases, and σ is the sigmoid function producing a probability of the positive class (5 kg load).

*3.6.2 ResNet 18 and ResNet 50*

ResNet architectures employ residual learning ((Eq. (3)), where each block predicts a residual mapping:

$$y = F\big(x, (\omega_i)\big) + x \qquad (3)$$

Here, *x* is the input to the block, *F* is the residual function (two convolution layers with ReLU), and *y* is the output. The context of deep residual networks enables efficient gradient propagation and mitigates the vanishing gradient problem in deep architectures, discussed by Taori and Lim [13]. Residual connections,

- Enable training of deeper networks without degradation.
- Preserve low-level features through skip connections.

*3.6.3 VGG19 Model*

VGG-19 has 5 Conv-blocks, each ending with 1 Max-Pooling layer, giving 5 Max-Pooling layers**.** Across these blocks, there are 16 Convolutional layers, and each includes a ReLU activation, totalling 16 ReLU layers. At the classifier end, there are 3 Fully Connected layers, where the final layer performs classification (Softmax applied here). Although it is not as computationally efficient as ResNet or DenseNet, it is employed as a standard for measuring performance on small EEG scalogram datasets.

*3.6.4 Custom Fine-Tuned CNN Model*

In this research, the customized residual Convolutional Neural Network (CNN) model with 13 convolutional layers is utilized. The model combines batch normalization and ReLU activation function to provide stable training and useful non-linear feature learning. It has six residual blocks that have two convolutional layers with identity skip-connections to achieve efficient gradient flow during backpropagation. The feature maps are gradually reduced in size by using three max-pooling layers. The convolutional layers are stacked in a hierarchy to obtain features at the different levels of abstraction. Spatial dimensionality is further compressed in the form of an adaptive average pooling layer before being classified. In order to facilitate generalization and reduce overfitting, dropout regularization is used six times across the network, both in the lower layers and the head of the classifier. The component of the classification comprises three layers, all of which are interconnected, and the third layer outputs one logit, which can be used in binary classification. Summarily, this model is a more profound residual learning-based CNN that incorporates efficient feature extraction with a stable dynamic training to attain precise binary image classification. Table 2 summarizes the main settings and parameters of the chosen deep learning models: 1D Convolutional Neural Network (1D-CNN), ResNet-18, ResNet-50, DenseNet-121, and VGG19.

**Table 2.** Representation of deep learning (DL) models' architecture and parameters.

| Deep Learning (DL) Model | Architecture | Key Parameters |
|---|---|---|
| 1D Convolutional Neural Network (1D-CNN) | Custom Sequential CNN | ➢ Max pooling layer: 3<br>➢ Residual block: 6<br>➢ ReLU layer: 15<br>➢ Dropout layers: 5<br>➢ Flatten layers: 1<br>➢ Output activation: Sigmoid<br>➢ Optimizer: Adam (lr = 0.0001, weight decay = 1e-4)<br>➢ Loss: Binary cross-entropy with logit<br>➢ Epochs: 50<br>➢ Batch Size: 16<br>➢ Early Stopping: patience = 5 |
| ResNet - 18 | Residual CNN | ➢ Max pooling layer: 1<br>➢ Residual block: 8<br>➢ ReLU layer: 17<br>➢ Dropout layers: 1<br>➢ Flatten layers: 1<br>➢ Output Activation: LogSoftmax<br>➢ Optimizer: Adam (lr = 0.005, weight decay = 1e-4)<br>➢ Loss: Negative log likelihood<br>➢ Epochs: 50<br>➢ Batch Size: 16 |
| ResNet - 50 | Deep Residual CNN | ➢ Max pooling layer: 1<br>➢ Residual block: 16<br>➢ ReLU layer: 49<br>➢ Dropout layers: 0<br>➢ Output Activation: Sigmoid<br>➢ Optimizer: Adam (lr = 0.0001, weight decay = 1e-5)<br>➢ Loss: Binary cross-entropy<br>➢ Epochs: 50<br>➢ Batch Size: 16<br>➢ Early Stopping: patience = 4<br>➢ Learning rate scheduler = StepLR (step size = 7, Gamma = 0.8) |
| DenseNet-121 | Densely Connected CNN | ➢ Max pooling layer: 1<br>➢ Dense block = 4<br>➢ Transition layer = 3<br>➢ ReLU layer: 120<br>➢ Dropout layers: 1<br>➢ Flatten layers: 0<br>➢ Output Activation: LogSoftmax<br>➢ Optimizer: Adam (lr = 0.0005, |

| | | |
|---|---|---|
| | | weight decay = 1e-4)<br>➢ Loss: Negative log likelihood<br>➢ Epochs: 50<br>➢ Batch Size: 16<br>➢ Learning rate scheduler = SterLR (step size = 5, Gamma = 0.8) |
| VGG19 | Deep Sequential CNN | ➢ Max pooling layer: 5<br>➢ Convolution layers = 16<br>➢ Fully connected layers = 3<br>➢ ReLU layer: 18<br>➢ Dropout layer: 1<br>➢ Flatten layer: 1<br>➢ Output Activation: LogSoftmax<br>➢ Optimizer: Adam (lr = 0.0005, weight decay = 1e-4)<br>➢ Loss: Negative log likelihood<br>➢ Epochs: 50<br>➢ Batch Size: 16<br>➢ Learning rate scheduler = SterLR(step size = 5, Gamma = 0.8) |

## 4 Model Performance Parameters

The performance of the model was measured after 50 epochs using conventional measures, such as training and validation accuracy, recall, and confusion matrices. During these periods, training and measures of validation were monitored to evaluate the learning behaviour and stability of the model. Accuracy was used as the primary metric to evaluate the overall classification performance, and recall provided information about the ability of the model to find relevant cases in the data. Also, the confusion matrix was used to study the distribution of correct and wrong classifications between various classes. These parameters of evaluation give together a clear knowledge of the behaviour and form of the model, which is the basis of assessing its performance in the given experimental conditions.

The number of parameters in a deep learning model indicates its ability to learn complex and meaningful features. Models with many parameters usually have higher representational power, but they also require a larger amount of training data and proper regularisation techniques to avoid overfitting. Deeper architectures, which contain more blocks and non-linear layers, learn features in multiple stages. This helps the model to capture hierarchical patterns, starting from low-level features and gradually moving to high-level representations. However, an increase in the number of parameters also results in higher computational cost, more memory requirements, and longer training time. In practical applications such as EEG-based image or time–frequency classification, models with a moderate number of parameters and well-designed connectivity (such as skip connections or dense connections) usually provide a better balance between accuracy and generalisation, especially when the available dataset is limited.

# 5 Result

This paper is an example of deep learning models that have been created to categorize the various. carrying activity at the shoulder during real-time gait movement with the use of EEG biomarkers. The data set comprised 8932 images, and it was tested with classifiers such as ResNet18, ResNet50, VGG19, DenseNet121 and Own CNN. The base rate of the EEG Scalogram, which was 8 electrodes, was used to compare the no-load and 5 kg load. The data was divided into three varying subsets: training (80%), validation (10%), and testing (10%). This rigorous evaluation provided a strong and thorough evaluation of the accuracy in classification based on different electrode combinations. The findings proved that even low-intensity physical activities cause significant alterations in cortical dynamics, especially frequency bands that are related to concentration, motor organization, and sensory organization.

Confidence Interval: To quantify the statistical reliability of the obtained classification accuracies, 95% confidence intervals (CI) were computed assuming a binomial distribution of correct and incorrect predictions. For each model and each experimental setting (combined electrodes and lobe-wise datasets), the confidence interval was calculated using the normal approximation of the binomial proportion as shown in Eq. (4):

$$CI_{95\%} = p \pm 1.96\sqrt{\frac{p(1-p)}{n}} \quad (4)$$

Where $p$ is the observed classification accuracy expressed as a proportion, and $n$ is the number of samples in the validation set.

## 5.1 Classifier Performance for Combined Electrodes

To enhance the spatial representation, the EEG data of all eight electrode sites were collected. The analysis was performed on (F1, F2, FC3, FC4, Fz, Cz, O1 and O2). This integrated dataset of the neural was able to have a closer look at cortical activity, as electrode readings provided a more descriptive approach to the network to identify spatially diffused patterns relating to a physical load. The analysis of the deep learning architecture of the joint electrode data showed that the proposed deep learning architecture was sufficiently able to distinguish the two loading conditions successfully using CWT-derived. EEG scalogram images. The application of the signals of the electrode sites provided a complete yield of spatial representation of brain activity during load-lifting tasks, which makes the model able to capture different brain activities associated with workload and involvement. The study aimed to determine the brain regions that showed the highest specific neural patterns of identifying differences between physical load during actual load-carrying tasks.

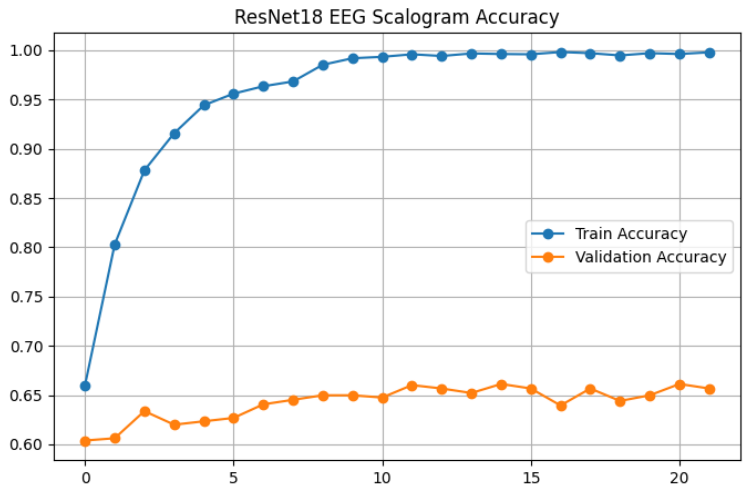
Three of the suggested models were used to assess the classification efficacy in brain regions. There are other measures of assessment, which were used: Precision, Recall, and F1-score. The F1-score is the harmonic average of Precision and Recall, so giving a fair evaluation of. model performance, particularly where there is a disparity in data distribution. A high F1 Score is an indication of a good balance between precision and recall, and a low F1 Score indicates a poor balance. The model is poorly performing.

Among the evaluated architectures, ResNet18 achieved the highest validation accuracy of 66.13% and a 95% confidence interval (CI) of 63.94–68.32, outperforming other transfer learning models. In contrast, DenseNet121 exhibited the lowest classification accuracy of 60.05% (95% CI: 57.78–62.32), while the custom CNN, VGG19, and ResNet50 achieved intermediate performance levels as shown in Table 3. These results indicate that the residual connections in ResNet18 enabled more effective feature extraction from the time–frequency representations, leading to improved classification of load-related EEG patterns.

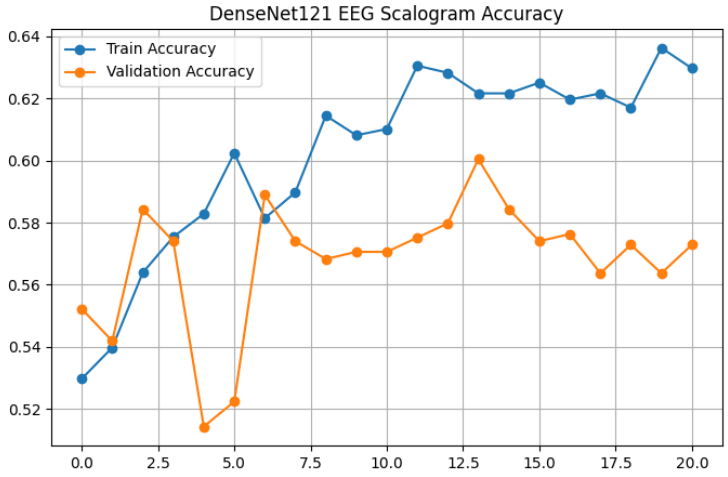
**Table 3**. Performance of classifier accuracy and Confidence Interval (CI) using various Deep Learning models for classifying different load conditions.

| CNN Model | Accuracy (%) | Confidence Interval (CI) |
|---|---|---|
| ResNet18 | 66.13 | 63.94 - 68.32 |
| ResNet50 | 62.46 | 60.22 - 64.70 |
| VGG19 | 64.52 | 62.30 - 66.74 |
| DenseNet121 | 60.05 | 57.78 - 62.32 |
| Own CNN | 61.06 | 58.80 - 63.32 |

Figure 7(A) illustrates the training and validation accuracy of the ResNet18 model applied to EEG scalogram data across multiple epochs for classifying various shoulder loads (No load and 5 kg Load) for the combined brain lobes. The training accuracy (blue line) increases and reaches 100 % accuracy, which shows Resnet 18 has learned the training data very well. However, the validation accuracy (orange line) is stable around 60-65%. Similarly, training accuracy (blue line) of the DenseNet121 model reaches a maximum 64% and validation accuracy (orange line) around 60.0% as shown in Figure 7(B). This indicates that DenseNet121 provides lower validation accuracy compared to other models. Out of five deep learning models, ResNet18 has high competence among other models, including our own CNN, with an accuracy of 65 %.




A

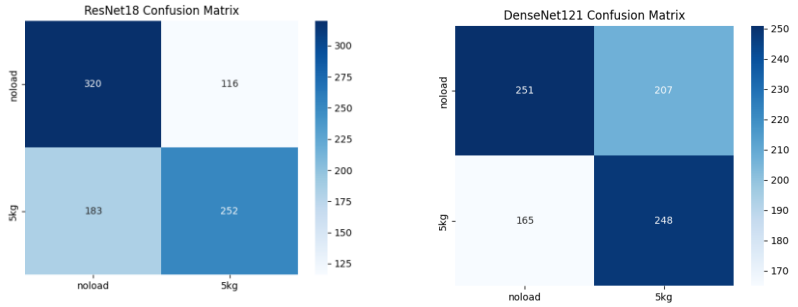



B

**Fig. 7**. Graphical representation of training and validation accuracy and loss curve for different epochs: (A). ResNet18, and (B). DenseNet121

The superior performance of ResNet18 can be attributed to its optimized architectural depth and incorporation of identity-based residual skip connections, which ensure efficient gradient backpropagation and alleviate the vanishing or exploding gradient problem commonly encountered in deep convolutional networks. Its moderate parameterization and shallower receptive hierarchy enable robust extraction of discriminative time–frequency feature embeddings from CWT-derived EEG scalograms for physical load classification, while maintaining a high bias–variance equilibrium across training and validation phases. The transfer learning model of DenseNet121 has dense connectivity designed to promote feature reuse, exhibits substantial feature redundancy and inter-layer correlation, which can lead to overfitting and reduced generalization when trained on limited, low-dimensional EEG datasets. Similarly, ResNet50, though equipped with residual blocks, introduces excessive computational overhead, gradient noise, and optimization instability due to its deeper topology and increased convolutional kernel density.

A confusion matrix is an important tool for evaluating the performance of a classification model. It demonstrates how well the model's outcomes match the actual values. For ResNet18, predicted 320 observations correctly and 118 observations are misclassified for the no load condition, with an accuracy of 73.39 %, as shown in Figure 8(A). Similarly, for a 5 kg load, 252 samples are correctly classified, and 183 observations are misclassified with an accuracy of 57.93%. DenseNet121 model, no load class 251 observations are correctly classified, and 207 samples are misclassified with an accuracy of 54.80 %. 5 kg load condition, 248 observations are correctly predicted, and 165 are misclassified with an accuracy of 60.04 %, as shown in Figure 8(B).




A B

**Fig. 8**. Confusion matrix for the combined electrode for classifier validation: (A). ResNet18, and (B). DenseNet121

## 5.2 Classifier Performance by Brain Lobel

To identify the most informative brain regions for distinguishing between different physical loads (no load and 5 kg load), three key lobes were analyzed: the frontal cortex (FP1, FP2, Fz), the Central (FC3, FC4, Cz) and the occipital (O1, O2). EEG Scalogram images from these regions were extracted and processed through Deep learning models to evaluate classification accuracy. This study aimed to identify the specific brain regions that exhibited the most distinctive neural signatures for distinguishing variations in physical load during actual locomotion.

In order to evaluate the classification efficacy of the suggested models for brain regions, three additional assessment metrics were utilized: Precision, Recall, and F1-score. The F1-score serves as the harmonic mean of Precision and Recall, thus offering an equitable assessment of model performance, especially in contexts where there is a disparity in data distribution. A high F1 Score suggests a good balance between precision and recall, and a low F1 score indicates the model has poor performance. Table 4 compares the performance of brain lobes with all five deep learning models for the F1 score. The occipital cortex region provides better classifier performance with a higher F1 score. For example, the ResNet18 achieved a 0.69 F1 score compared to the remaining brain lobes. The frontal brain region F1 score was lower for all models when compared to other brain regions.

**Table 4**. Comparison of the F1 score for all five deep learning models across three brain regions for classifying no load and 5 kg load

| **Brain Region** | **Model** | | | | |
|---|---|---|---|---|---|
| | **ResNet 18** | **ResNet50** | **VGG19** | **DenseNet121** | **Custom CNN** |
| Frontal cortex (FP1, FP2, Fz) | 0.535 | 0.595 | 0.55 | 0.565 | 0.394 |
| Central (FC3, FC4, Cz) | 0.585 | 0.605 | 0.585 | 0.64 | 0.57 |
| Occipital (O1, O2) | 0.69 | 0.67 | 0.65 | 0.585 | 0.545 |

Precision indicates how many predicted samples as positive are actually positive, reflecting the model's reliability in minimizing false positives. Recall, on the other hand, quantifies the proportion of correctly predicted positive samples among all actual positives, indicating the capability of the model to detect relevant instances and reduce false negatives. In our study, the occipital region of the brain with the ResNet18 model has high precision and recall values, such as 0.695 and 0.69, when compared with other brain regions with other models, as shown in Table 5.

**Table 5**. Representation of precision, recall for all five deep learning models across three brain regions for classifying no load and 5 kg load

| **Brain Region** | **Model** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **ResNet 18** | | **ResNet50** | | **VGG19** | | **DenseNet121** | | **Custom CNN** | |
| | **Precision** | **Recall** | **Precision** | **Recall** | **Precision** | **Recall** | **Precision** | **Recall** | **Precision** | **Recall** |
| Frontal cortex (FP1, FP2, Fz) | 0.595 | 0.565 | 0.605 | 0.6 | 0.575 | 0.565 | 0.565 | 0.565 | 0.515 | 0.505 |
| Central (FC3, FC4, Cz) | 0.59 | 0.585 | 0.605 | 0.605 | 0.585 | 0.585 | 0.645 | 0.64 | 0.635 | 0.595 |
| Occipital (O1, O2) | 0.695 | 0.69 | 0.675 | 0.675 | 0.605 | 0.605 | 0.59 | 0.585 | 0.635 | 0.585 |

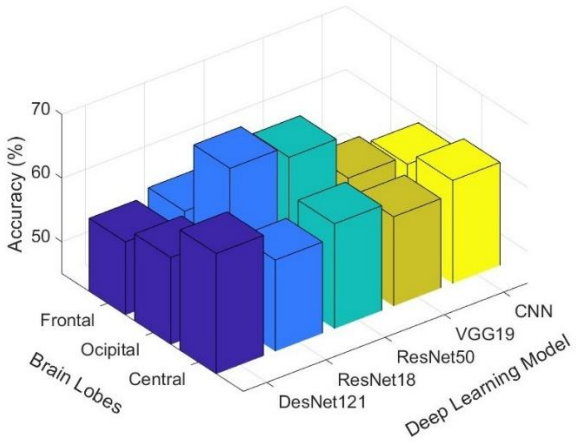



**Fig. 9.** Classification accuracy of five deep learning models with brain lobes using EEG biomarkers.

As shown in Figure 9, the ResNet18 model outperformed the other classifiers, achieving 69.09 % accuracy with the occipital cortex brain region, 67.27 % for ResNet50 with the same occipital lobes. The DenseNet121 model also performed well, reached 63.91 % in the central cortex of brain lobes, 61.47 % for ResNet50 and 61.16 % for our own CNN in the same central lobe of the brain. In contrast, the Frontal region of the brain showed lower

classifier performance with accuracy below 57.1 % for all five deep learning models. These results demonstrate that the occipital cortex with the ResNet18 model significantly enhances classification accuracy and is most effective in distinguishing between different physical load conditions during real gait movements. It means that the visual cortex produces distinct EEG patterns associated with. cognitive functioning as one indulges in weightlifting. The enhanced efficacy of the residual nature of the learning structure can be attributed to the fact that ResNet18 facilitates the more efficient flow of gradients and retains the necessary representations of EEG signals without any loss, succumbing to overfitting.

Table 6 shows the confidence interval analysis across three cortical regions of classification capability among the evaluated deep learning architectures for discriminating between no-load and 5 kg load conditions. In the frontal cortex VGG19 model stronger performance (56.1–63.8) compartely with other DL models, and the custom CNN exhibit the lowest discriminative stability (45.0–52.8). In the central cortex, ResNet18 achieves the widest and highest interval (57.6–70.2), and the remaining models show overlapping but slightly lower ranges. For the occipital cortex region, ResNet50 and VGG19 produce the highest confidence ranges (63.1–75.1 and 61.1–73.9, respectively) when compared with other DL models.

**Table 6**. Analysis of confidence interval (CI) for all five deep learning models across three brain regions for classifying no load and 5 kg load

| DL Model | Frontal Cortex | Central Cortex | Occipital Cortex | Inferences |
|---|---|---|---|---|
| ResNet18 | 52.6 - 60.4 | 57.6 - 70.2 | 52.2 - 65.1 | ➢ Central cortex - ResNet18 has superior feature extraction from sensorimotor activity<br>➢ Strong CI in occipital region - spatially specific neural dynamics influence deep network performance in load classification tasks |
| ResNet50 | 53.9 - 61.6 | 52.9 - 65.7 | 63.1 - 75.1 | |
| VGG19 | 56.1 - 63.8 | 55.1 - 67.8 | 61.1 - 73.9 | |
| DenseNet121 | 52.6 - 60.3 | 52.6 - 65.5 | 54.0 - 66.9 | |
| Own CNN | 45.0 - 52.8 | 54.8 - 67.5 | 52.6 - 65.1 | |

## 6 Discussion

Load-lifting activities are a complex combination of muscular work and neural work. synchronization, which leads to distinct cortical activation patterns which can be effectively. monitored by electroencephalography (EEG). The current study formed a framework. a deep learning-based approach to classify load lifting scenarios continuously. EEG scalogram representations based on wavelet transform (CWT). The objective of the investigation was to determine how different spatial electrode arrangements and neural activity affected each other. designs of classification efficacy. The built-in electrode test gave an in-depth illustration of the cortical activity, which enables a specific differentiation between. no-load conditions and 5 kg load conditions. This arrangement achieved validation. The accuracy of 66.13% obtained with the use of the ResNet18 architecture

demonstrates the advantage of. combining the activity of different parts of the brain to record scattered neural activity. during physical tasks.

Further examination of electrodes, which were specific to various brain regions, proved that the occipital lobe had the best accuracy of 69.09, which highlights the importance of the role of the visual and perceptual processing in controlling motor coordination in load-lifting activities. These findings indicate that the occipital region activity is sufficiently powerful in task differentiation, perhaps because of improved visual attention and geographical awareness. There is also the higher effectiveness of the ResNet18. The efficiency of the model compared to other frameworks of transfer learning is highlighted by the residual learning of complex temporal-spatial patterns that occur in the EEG biomarkers.

Kumar et al. [19] relied on a hybrid system that combines CNN and LSTM in recognition. using EEG indicators of physical fatigue, the maximum classification accuracy was achieved 61.2%. Though their hybrid method used dynamics over time, they centred on 1D signal properties during preprocessing, limiting the ability of the model to detect complex visual frequency. correlations. By contrast, the feature representation used in our research was a scalogram-based one convolutional network that has the ability to learn both frequency-domain and spatial relationships. more effectively. Besides, the residual architecture of the ResNet18 model was improved. propagation of features, which leads to better discriminative abilities and generalization. across different subjects. In this way, the strategy that we suggest is more efficient and specific. technique of EEG patterns classification of cognitive and load-related patterns.

Table 7 offers a comparative study of dominant methodologies and the proposed one. technique, showing that the latter has improved performance in relation to accuracy. assessment metrics. Conversely, in most of the existing studies, one or the other is predominant: unprocessed EEG measurements, manually engineered temporal features, a mixture of bio signals, such as EMG and data provided by wrist-band sensors, and the accuracy has been measured to be reported as normally coming below, from 60% to 66%. Although these techniques are sufficiently effective to identify overall workload trends, they often fail in their ability to detect fine-scale spatial-frequency dynamics over various cortical areas.

**Table 7**. Evaluation of the proposed model against current leading studies in EEG-based load or cognitive workload classification.

| Article (Year) | Purpose | Algorithms / Models used | Highest Accuracy (%) |
|---|---|---|---|
| Rebsamen et al. [11] | Detect cognitive workload levels from EEG during arithmetic tasks | Spectral features + SMLR classifier | 62.0 |
| Borisov et al. [16] | Detect workload from multimodal wearable signals | Feature-based ML/ensemble models | ≈63.0 |
| **Our Approach: Deep learning + Brain lobes** | **Classification of load-lifting conditions (no-load vs 5 kg) using EEG scalograms.** | **Transfer learning models — ResNet18, ResNet50, VGG19, DenseNet121 — and a custom CNN** | **66.13 (All-electrodes, ResNet18); 69.09 (Occipital region, ResNet18)** |

## 7 Conclusion and Future Work

This experiment revealed that EEG data converted to CWT-based scalograms, which are processed by deep learning models, are effective in categorization of physical load conditions during walking. ResNet18 was found to be the best in the evaluated architectures, which confirms the ability of the neural architecture to retrieve meaningful neural features that are related to load-related cortical activity. The occipital part of the brain performed above the average, and this means that it is very vital in processing the neural responses caused by load. These results indicate the opportunities of the EEG-based deep learning models in objective physical loading monitoring and adaptive human-machine interaction systems.

Further research will be done to increase the size of the dataset, the number of levels of the loads and the size of the population of the subjects to enhance generalization and strength. Further, a multimodal combination of EEG and electromyography (EMG) and inertial measurement units (IMUs), to measure body motion and body orientation, will be investigated to improve classification accuracy. Embodied systems will also be explored using real-time systems and advanced architectures like attention-based and transformer models to be deployed in robotic exoskeletons, rehabilitation applications, and brain-computer interface applications.

## References

1. J. S. Boschman, H. F. van der Molen, J. K. Sluiter, M. H. Frings-Dresen, Musculoskeletal disorders among construction workers: a one-year follow-up study. BMC Musculoskelet. Disord. **13**, 196 (2012). https://doi.org/10.1186/1471-2474-13-196
2. A. M. Brouwer, M. A. Hogervorst, J. B. van Erp, T. Heffelaar, P. H. Zimmerman, R. Oostenveld, estimating workload using EEG spectral power and ERPs in the n-back task. J. Neural Eng. **9**, 045008 (2012). https://doi.org/10.1088/1741-2560/9/4/045008
3. A. Craik, Y. He, J. L. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: a review. J. Neural Eng. **16**, 031001 (2019). https://doi.org/10.1088/1741-2552/ab0ab5
4. M. Dai, D. Zheng, R. Na, S. Wang, S. Zhang, EEG classification of motor imagery using a novel deep learning framework. Sensors **19**, 551 (2019). https://doi.org/10.3390/s19030551
5. M. Goršič, B. Dai, D. Novak, Load position and weight classification during carrying gait using wearable inertial and electromyographic sensors. Sensors **20**, 4963 (2020). https://doi.org/10.3390/s20174963
6. F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, B. Arnaldi, A review of classification algorithms for EEG-based brain-computer interfaces. J. Neural Eng. **4,** R1–R13 (2007). https://doi.org/10.1088/1741-2560/4/2/R01
7. M. Aharon, M. Elad, A. Bruckstein, Learning sparse representations of high dimensional data on large-scale dictionaries. Adv. Neural Inf. Process. Syst. **23** (2010).
8. M. Mahmud, M. S. Kaiser, A. Hussain, S. Vassanelli, Applications of deep learning and reinforcement learning to biological data. IEEE Trans. Neural Netw. Learn. Syst. **29**, 2063–2079 (2018). https://doi.org/10.1109/TNNLS.2018.2790388
9. M. I. M. Refai, A. Moya-Esteban, M. Sartori, Electromyography-driven musculoskeletal models with time-varying fatigue dynamics improve lumbosacral joint moments during lifting. J. Biomech. **164**, 111987 (2024). https://doi.org/10.1016/j.jbiomech.2024.111987

10. M. Z. Ramadan, A. M. Ghaleb, A. E. Ragab, Using electroencephalography (EEG) power responses to investigate the effects of ambient oxygen content, safety shoe type, and lifting frequency on the worker's activities. Biomed Res. Int. 2020, 7956037 (2020). https://doi.org/10.1155/2020/7956037
11. B. Rebsamen, K. Kwok, T. B. Penney, EEG-based measure of cognitive workload during a mental arithmetic task. Commun. Comput. Inf. Sci. **174**, 62 (2011). https://doi.org/10.1007/978-3-642-22095-1_62
12. Z. Ren, R. Li, B. Chen, H. Zhang, Y. Ma, C. Wang, Y. Lin, Y. Zhang, EEG-based driving fatigue detection using a two-level learning hierarchy radial basis function. Front. Neurorobot. **15**, 618408 (2021). https://doi.org/10.3389/fnbot.2021.618408
13. S. Taori, S. Lim, Use of a wearable electromyography armband to detect lift-lower tasks and classify hand loads. Appl. Ergon. **119**, 104285 (2024). https://doi.org/10.1016/j.apergo.2024.104285
14. Y. R. Tabar, U. Halici, A novel deep learning approach for classification of EEG motor imagery signals. J. Neural Eng. **14**, 016003 (2017). https://doi.org/10.1088/1741-2560/14/1/016003
15. D. Totah, L. Ojeda, D. D. Johnson, D. Gates, E. M. Provost, K. Barton, Low-back electromyography (EMG) data-driven load classification for dynamic lifting tasks. PLoS ONE 13, e0192938 (2018). https://doi.org/10.1371/journal.pone.0192938
16. V. Borisov, E. Kasneci, G. Kasneci, Robust cognitive load detection from wrist-band sensors. Comput. Hum. Behav. Rep. **4**, 100116 (2021). https://doi.org/10.1016/j.chbr.2021.100116
17. X. Zheng, W. Chen, M. Li, T. Zhang, Y. You, Y. Jiang, Decoding human brain activity with deep learning. Biomed. Signal Process. Control **56**, 101730 (2020). https://doi.org/10.1016/j.bspc.2019.101730
18. A. Delorme, T. J. Sejnowski, S. Makeig, Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage **34**, 1443–1449 (2007). https://doi.org/10.1016/j.neuroimage.2006.11.004
19. Y. Zhang, Y. Wang, J. Jin, X. Wang, Sparse Bayesian learning for obtaining sparsity of EEG frequency bands based feature vectors in motor imagery classification. Int. J. Neural Syst. **27**, 1650032 (2017). https://doi.org/10.1142/S0129065716500325
20. G. Zhou, M. L. Lu, D. Yu, Tactile gloves predict load weight during lifting with deep neural networks. IEEE Sens. J. **23**, 18798–18809 (2023). https://doi.org/10.1109/JSEN.2023.3289670