Muscle-Guided Guitar Pedalboard: Exploring Interaction Strategies Through Surface Electromyography and Deep Learning

Davide Lionetti
Dept. of Art and Science
University of Ghent
Ghent, Belgium
davide.lionetti@ugent.be

Paolo Belluco
LWT3
Milan, Italy
paolo.belluco@lwt3.com

Massimiliano Zanoni
Dept. of Electronics, Information and Bioengineering
Polytechnic University of Milan
Milan, Italy
massimiliano.zanoni@polimi.it

Luca Turchet
Dept. of Information Engineering and Computer Science
University of Trento
Trento, Italy
luca.turchet@unitn.it

ABSTRACT

This paper explores a method to innovate the conventional interaction with a guitar pedalboard. By analyzing muscular contractions tracked via surface Electromyography (sEMG) wearable sensors, we aimed to investigate how to dynamically track guitarists' sonic intentions to automatically control the guitar sound. Two Recurrent Neural Networks based on Bidirectional Long-Short Term Memory were developed to analyze sEMG signals in real-time. The system was designed as a digital musical instrument that calibrates itself to each user during an initial training process. During training musicians provide their gestural vocabulary, associating each gesture to a corresponding pedalboard preset. The selection of the most effective features, in synergy with the best set of muscles, was conducted to optimize the learning rate of the system. The system was assessed with a user study encompassing seven expert guitar players. Results showed that, on average, participants appreciated the concept underlying the system and deemed it to be able to foster their creativity.

Author Keywords

Gesture To Sound Mapping, Guitar Pedalboard, Surface Electromyography, Deep learning, Recurrent Neural Network, Wearable Device.

CCS Concepts

•Applied computing → Sound and music computing; •Human-centered computing → Gestural input; *User studies;*

NIME'24, 4–6 September, Utrecht, The Netherlands.

1. INTRODUCTION

The swift progression of wearable technology alongside advancements in machine learning (ML) methodologies has opened avenues for innovative interaction paradigms within live music contexts, centering around the real-time detection of performers' gestures. This domain has strong ties with the embodied music cognition theory, which explores the connections between physical actions and musical expression [22]. Embodied engagement with music is a key element of musical experience, and the gestural properties of musical sound have been studied from multiple disciplinary perspectives, including Human-Computer Interaction (HCI) and the cognitive sciences [14, 37].

Digital musical instruments (DMIs) are based on the creation of action–sound mappings, in which the relationships between the physical energy of the input action may not necessarily correspond to that of the output sound [7]. In this scenario, gestural interaction design is a non-trivial problem, and ML techniques are increasingly utilized by researchers and artists to tackle its complexity in various ways [37, 32].

ML methods have been leveraged for the real-time analysis of electromyographic (EMG) signals, which are a source of meaningful information about the user's sonic intentions [30]. EMG is a biometric signal that represents muscle activity. It has been used in the biological and HCI domains as a highly sensitive method of capturing human movement. Lately, it has also been found to be an appropriate medium for sensing musical gesture.[39, 31]. The Surface EMG (sEMG) is a technique that allows non-invasive detection of muscle activity, without the use of needles (typically used in EMG medical applications) but instead using passive electrodes placed on the skin. Several scholars have utilized such a technique for the creation of a variety of DMIs [25, 36, 8, 19, 12].

In this paper, we investigate the relation between guitarists' sonic intentions and their muscular contractions, to ease the interaction with sound effects without the use of pedals. We present a proof of concept DMI which adapts the sound of the guitar based on the muscular activation of the musician[1]. We also describe a user study with seven

[1]A video of the prototye in use is available at `https://youtu.be/8_c5QavFDUA`

expert guitar players, which aimed to assess the following research question: *Are muscle contractions useful to musicians, as a new form of interaction with a pedal-board, to encourage exploration and foster creativity?*

Electric guitar pedals are commonly seen as a regular part of the instrument and not as an extension or enhancement of its characteristics. This is because the effects used in electric guitars usually have one-dimensional control possibilities, like the Wah-Wah pedal, or only offer the option to activate or deactivate some previously configured processes [10, 24]. Often, guitarists are distracted by the need to press multiple pedals while playing, to shape the guitar's sound. This motivated our research in exploring a strategy for interacting with sound effects that is different from the conventional one.

2. RELATED WORKS

Performing robust and accurate EMG signal decoding represents a critical challenge due to its noisy characteristics. Various types of research have addressed this problem exploiting Deep Learning techniques such as: hand gesture classification [21], [23], silent speech recognition [20], [18], stroke rehabilitation [26] and robot arm control [1] (for a review see Buongiorno et al. [6]).

Within the NIME community, various studies have investigated the role of gesture representation in sound-action mapping. Various scholars have investigated the use of sEMG signals to track users' gestures via wearable devices that measure bodily behaviors to manipulate digital media [28, 33]. Most of previous researches are based on forearm muscles, using the Myo armband device [25, 36, 8, 19, 12], or other custom-made sensor boards [9].

A relevant example is the *air guitar* [11], a DMI leveraging the guitarists' gestures performed in the air to control a guitar sound emulation engine, using two Myo armbands. The authors evaluate several excitation categories (excitation is a phase where there is energy transfer between a music performer and a musical object) to investigate how action-sound gesture coupling can be used to create a DMI not relying on a physical controller.

3. METHODS AND MATERIALS

3.1 Data acquisition

To integrate sEMG signals into artistic practices reliably and robustly, we used the wearable device developed by LWT3[2] (see Fig. 1). This wearable computer system enables the tracking of biometric signals from multiple muscle areas, with a sampling rate of 1000Hz and 22-bit resolution ADC for 8 channels in a double differential configuration. The acquired data is then transmitted to the platform analysis via a wired USB 2.0 communication channel. During the user study described in Section 5, electrode pads were placed by following the protocols of the *Atlas of Muscle Innervation Zone* [3], to minimize cross-talk and enhance the acquisition quality.

For the acquisition and storage of sEMG signals from the sensor board we leveraged the LWT3's proprietary data processing application, which also computes various features and provides real-time visualization of muscular activities (see Fig. 2). To further enhance data quality, we implemented an additional *pre-filtering stage* to mitigate movement artifacts and electromagnetic interferences during live performances. This stage incorporates a fifth-order Butterworth bandpass filter (30-300 Hz) to attenuate high-frequency motion artifacts and a harmonic band-stop Notch filter (centered at 50 Hz) to suppress power line noise.

[2]https://lwt3.com/

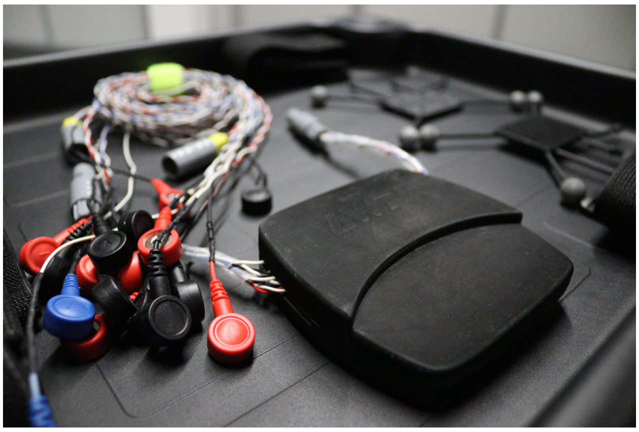


Figure 1: Data acquisition platform: 8-channels sEMG wearable sensor board. Two sides, four channels for each side.

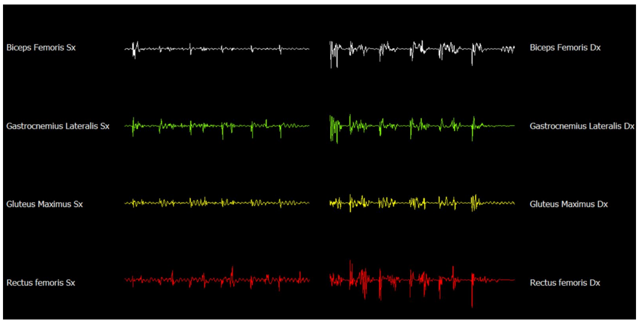


Figure 2: Muscular contraction live plots from the data elaboration application.

3.2 Muscle Selection Process

The selection of appropriate muscles is paramount in ensuring accurate gesture classification, as distinct guitar techniques elicit unique patterns of muscle activation. To establish a robust muscle selection criterion, we conducted multiple 30-second acquisitions for each gesture. Then we selected two low-level features in the temporal domain (that required lower computational power), from those most widely used in the literature [29]:

1. **Root Mean Square (RMS):** is related to the constant force and non-fatiguing contraction, it reflects power activation and is directly proportional to the exerted force. It relates to standard deviation, which can be expressed as

$$RMS = \sqrt{\frac{\sum_{k=1}^{N} x_k^2}{N}}. \tag{1}$$

where N denotes the length of the signal and x_n represents the EMG signal in a segment n.

2. **Zero crossing (ZCR):** is the number of times that EMG signals cross zero in a window of length N. The threshold value is 20 mV. It can be formulated as:

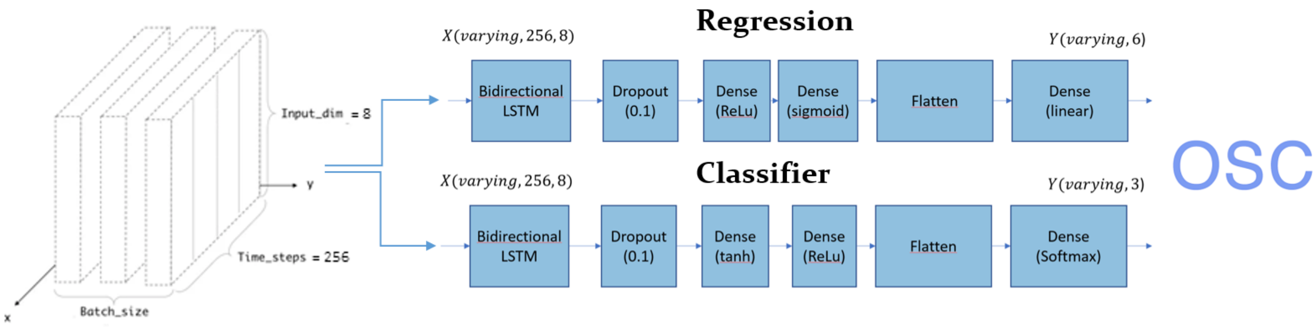


Figure 3: The models' architecture, specifying input and output shapes and activation functions.

$$ZCR = \frac{1}{2(N-1)} \sum_{i=1}^{N-1} \begin{cases} 1 & \text{if } x(i) \cdot x(i+1) < 0 \\ & \text{and } |x(i) - x(i+1)| > \text{thr.} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

ZCR is related to slope sign change (SSC) and it gives a rough estimation in the frequency domain.

After that two consecutive stages were performed: In the first, the ZCR was used to assess the level of activation over time of each muscle, during the unfolding of gestures. We discarded the muscles with ZCR values that fell below a predefined threshold. Then in the second stage, groups of eight muscles were evaluated together, selecting the group with the highest RMS average value.

Through extensive testing, we observed that in a resting position (with our configuration test), the ZCR values ranged from 40 to 70, while the RMS was around 2.3 mV compared to the baseline. These observed ranges serve to define our heuristic activation threshold: 50 slop change (ZCR) and 4.6 mV (RMS) into a 500 ms window.

In table 2 the muscles are ranked based on their mean ZCR and RMS values, we discarded those below the thresholds (i.e., the last six). Muscles with consistently low RMS and ZCR values were excluded from further consideration. To refine the muscle selection process and identify muscle activation patterns associated with each guitar gesture, the ZCR was utilized to evaluate average activation levels within predefined time intervals. By combining the RMS and ZCR features, we determined the optimal set of eight upper limb muscles for accurately classifying each specific gesture:

3.3 ANNs Design

We applied a supervised learning approach, by developing two Recurrent Neural Networks (RNN) that work in parallel: one **gesture classifier** and a **regression network**. The first classifies the performed guitar techniques dynamically setting the pedalboard's sound preset. The other is trained for regression tasks by providing examples of inputs associated with desired sound effects parameters' outputs. Once the model is trained, the user performs by moving between (and beyond) the example positions to shape dynamically the sounds via gestures. The models were implemented using TensorFlow (version 2.12.0) and the Keras API.

3.3.1 Dataset creation

The system was developed to be **intra-user** (i.e., customized for each user), to model the nuances and habits of each guitarist distinctly. To achieve this, a systematic procedure was employed for training each instance of the system. Seven distinct guitar techniques were chosen: *fingerpicking, strumming, bending, down picking, alternate picking, tapping, and pull-off/hammer-on.* Each technique was paired with a corresponding guitar riff, serving as a representative exemplar of its characteristic motion. This pairing ensured uniformity in movements across diverse participants. The training process was carried out by following these sequential steps:

1. Users were instructed to perform the seven pre-selected riffs. Dual data acquisitions were conducted for each technique, prompting the subject to play at opposing levels of muscle contraction. This approach aimed to capture the maximum and minimum values, essential for training the regression model tasked with discerning the different levels of contraction exerted during the execution of the gestures.

2. Each acquisition was executed using an electric guitar, with participants adopting a stationary standing position. The duration of each session was fixed at 60 seconds, maintaining a consistent tempo of 120 beats per minute (BPM).

3. Subsequently, a dedicated function was employed to export a filtered RMS version of each acquisition. This process facilitates the preservation of maximum and minimum values for each muscle, enabling subsequent normalization within the $[0, 1]$ range.

The signals were captured from the eight selected muscles presented in Section 4.1 using the sensor board presented in Section 3.1.

Each user's dataset was normalized between $[0, 1]$ with a Mix-Max strategy, to ease the inferences extraction by promoting fair treatment of features, facilitating faster convergence, and enhancing the network's ability to generalize its learned knowledge. The dataset is publicly available, intending to address the lack of public sEMG datasets for guitar techniques [3].

3.3.2 Models' Architecture

Figure 3 depicts the two models' architecture, specifying the input/out shapes and the activation functions of each layer. Both architectures are composed of a Bidirectional Long Short-Term Memory (BLSTM) input layer followed by two fully connected dense layers. Notably, it has been demonstrated that incorporating dense layers in conjunction with the recurrent layer can yield superior outcomes compared to models solely relying on recurrent layers [16].

[3]https://github.com/EllDy96/Augmented-Guitar-Pedalboard

Hyperparameter	Value
Input Time Steps	256
Regularization Strategy	L1(factor= 0.001)
Dropout Rate	0.1
Optimization Algorithm	Adam
Learning Rate	0.0001
Loss Function	Categorical Cross Entropy
Batch Size	32
Early Stopping	Monitor 'loss function' with patience 10
Early Stopping	Monitor 'validation loss' with patience 10

Table 1: Models' hyperparameter values.

The BLSTM layer, which incorporates an RNN with LSTM cells, is known for well capturing temporal patterns in gesture recognition tasks using sEMG signals [34]. The bidirectionality of the BLSTM allows it to capture temporal dependencies in both the past and future directions, enhancing the model's predictive capabilities [2]. To enhance the generalization and feature extraction capabilities of the RNN, two Dense layers were added after the BLSTM layer. The models also included Dropout layers, which randomly drop out a certain proportion of neurons during training to prevent over-fitting.

The two used cost functions were categorical cross-entropy for classification and Mean Square Error for regression. The weights of the models were updated using the ADAM optimization algorithm, a combination of gradient descent with momentum and RMS propagation [17]. The learning rate, which controls the magnitude of weight updates during training, was manually selected to ensure optimal convergence and stability. Additionally, L1 kernel regularization was applied to the loss function to prevent over-fitting by encouraging sparse weight values in the network. The source code and scripts are openly accessible on Google Colab ??.

3.3.3 *Hyperparameter Tuning*

We performed an intensive manual search to strike a balance between model complexity and generalization capability. The results are shown in Table 1 (both for classification and regression). The goal was to define a lightweight model suitable for real-time applications to be integrated into wearable devices. We started from 2 million trainable parameters and were able to reduce them to 31,971 for the classifier (450 Kb storage space) and 56,550 for the regression model (735 Kb). The main factor that enabled this sharp reduction in parameters, while maintaining high performance, was the change in architecture from standard to bidirectional LSTM.

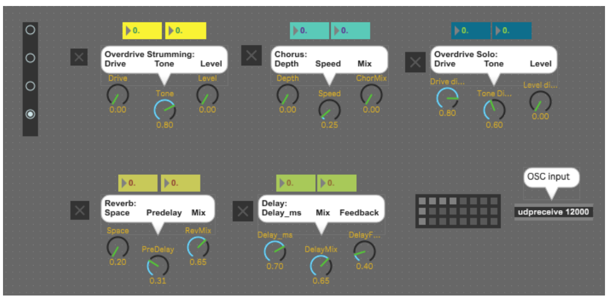


Figure 4: Max/Msp patch for sound parameters control.

3.4 Protocol Pipeline

The following steps compose the protocol pipeline (see Fig. 5) for acquiring, processing, and classifying sEMG signals in real-time to control the guitar sound:

Prefiltering Stage: A prefiltering stage is applied to ensure signal fidelity. This stage consists of an anti-aliasing filter and a high-pass filter (with a cutoff frequency of 5Hz) applied in cascade.

Data Acquisition and Processing: The wearable board transmits the signals to a computer system running a proprietary data acquisition platform. This platform performs crucial operations such as feature extraction and signal packaging. The packaged data is then simultaneously fed into two RNNs.

Gesture Classification: The first RNN is dedicated to gesture classification, selecting the pedalboard's preset based on the recognized gestures.

Effect Modulation: The second RNN operates as a regression model, enabling continuous modulation of the chosen effects based on the amount of muscular contraction.

Max/Msp 8 Patch: The predicted parameters from the RNN models are sent to a Max/MSP patch using the Open Sound Control (OSC) network protocol. The patch encapsulates a series of five Virtual Sound Technology (VST) plugins arranged sequentially to achieve the desired audio effects and modifications.

Audio Output and Processing: The path audio output is routed to a PA system.

3.5 Sound Processing Unit

The sound synthesis engine was developed in Max/MSP and relied on five imported VST plugins:

- **Overdrive**: *Overdrive TSC 1.1* developed by *Mercuriall* with three knobs to control: drive, tone, level.
- **Distortion**: *Distortion Greed smasher* developed by *Mercuriall* with three knobs to control: drive, tone, level.
- **Chorus**: *Chorus WS-1* developed by *Mercuriall* with three knobs to control: depth, speed, mix.
- **Reverb**: *Pro-R* developed by Fab Filter with three knobs to control: space, pre-delay (ms), mix.
- **Delay**: *Valhalla Super Massive Delay* developed by ValhallaDSP with three knobs to control: delay time (ms), mix, feedback.

We imported these five VSTs with `vst object`, connecting them with the data elaboration platform via OSC messages [38]. The patch receives the output of the two RNN models through two OSC messages: one for setting the sonic

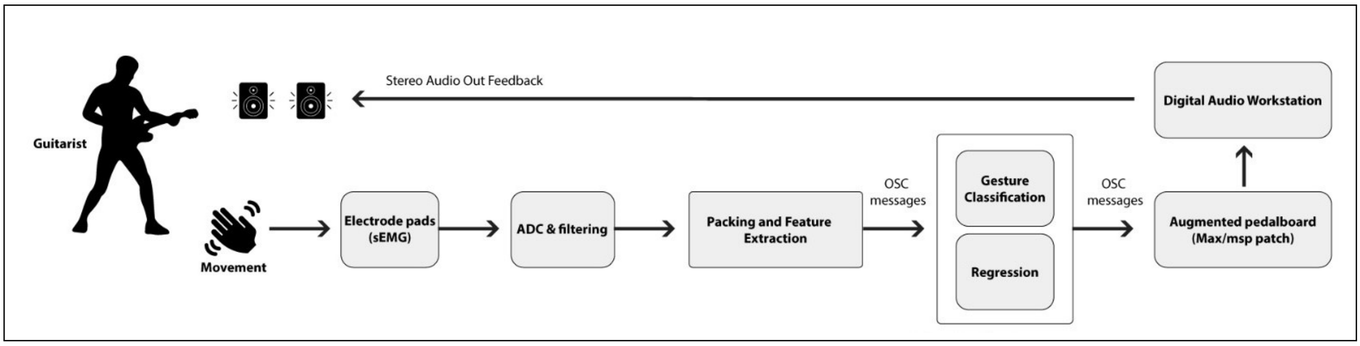


Figure 5: This image shows the entire signal flow, from its generation and acquisition from the user's body to the modification of the sound by the Max/Msp patch.

preset, the other containing a list of float in a $[0, 1]$ range for the real-time parameters modulation.

As it is possible to notice from Fig. 4, the patch has three knobs for each VST, allowing the modification of the effect's main parameters. These knobs are directly controlled by OSC messages allowing the regression network to change the pedalboard's parameters. The artist can save and recall sonic presets using the `preset object` and activate/deactivate each VST with the `toggle button`.

4. EXPERIMENTAL SET UP

The experimentation unfolded in three distinct phases, each incorporating multiple muscle acquisitions carried out with the wearable sensors boards described in Section 3.1:

1. **Muscle Selection**
   - *Objective:* To establish systematic criteria for muscle selection and assess the optimal upper limb muscles for accurate guitar gesture classification;
   - *Result:* The identification of eight specific arm muscles that significantly enhance classification outcomes (see Sections 3.2 and 4.1).
2. **Mapping Strategy**
   - *Objective:* To define the mapping of gestures to sound;
   - *Result:* The conceptualization and implementation of two RNNs architectures, encompassing the training process, dataset, and Max/MSP patch (see Sections 3.3.1, 3.3.2 and 3.5).
3. **Evaluation**
   - *Objective:* To assess the system's performance and effectiveness across multiple guitarists;
   - *Result:* Execution of a validation questionnaire to gauge the system's efficacy (see Section 5) together with the computation of the objective performance (Section 4.3).

In the following sections we present the results of each of these steps.

Figure 6: A picture of a participant during the acquisition session with electrode pads placed on arms.

4.1 Selected Muscles

Following the selection strategy described in Section 3.2, we performed an evaluation experiment to select the eight most relevant upper limb muscles for guitar technique classification. Twelve upper limb muscles were analyzed (see Fig. 7): *[Left Flexor Carpi Radiali, Left Extensor Carpi Radialis, Left Bicep Brachii Short Head, Left Brachioradialis, Right Flexor Carpi Radialis, Right Extensor Carpi Radialis, Right Bicep Brachii Short Head, Right Anterior Deltoid, Left Anterior Deltoid, Right Triceps, Left Triceps, Right hand, Left hand, Right Brachioradialis].* Among those, the following revealed to be the best for our purpose:

1. **Three forearm muscles**: including the *flexor carpi radialis the extensor carpi radialis* and *brachioradialis*, they are responsible for the wrist movement during guitar playing, which is particularly activated during techniques such as strumming, arpeggio, tremolo picking, alternate picking and down picking.

Muscles \ Gestures	Tapping		Arpeggio		Bending		Strumming		Pull-Off		Down pick		Alternate pick	
	ZCR	RMS	ZCR	RMS	ZCR	RMS	ZCR	RMS	ZCR	RMS	ZCR	RMS	ZCR	RMS
Left Flexor C. Radialis	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓
Left Extensor C. Radialis	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓
Left Bicep B, Short Head					✓	✓					✓	✓		
Left Brachioradialis	✓	✓	✓	✓			✓	✓			✓	✓	✓	✓
Right Flexor C. Radialis	✓	✓	✓	✓			✓	✓			✓	✓	✓	✓
Right Extensor C. Radialis	✓	✓	✓	✓			✓	✓			✓	✓	✓	✓
Right Bicep B. Short Head					✓	✓					✓	✓		
Right Anterior Deltoid							✓	✓			✓	✓		
Left Anterior Deltoid														
Right Triceps														
Left Triceps														
Right hand														
Left hand														
Right Brachioradialis														

Table 2: This table highlights which is the best muscle to classify a specific guitar gesture based on ZCR and RMS mean values for a standard right-hand guitarist.

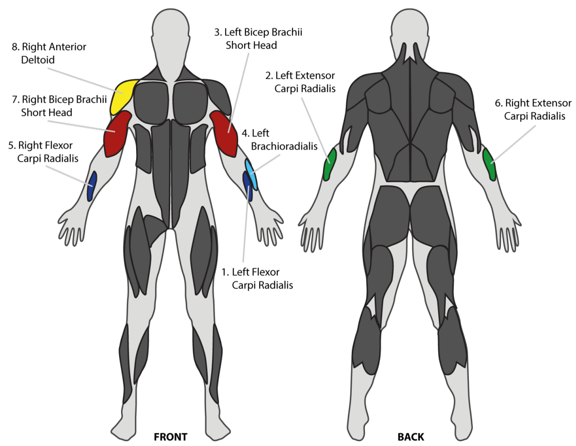


Figure 7: The eight selected muscles to classify guitar gestures.

2. **Right anterior Deltoid**: This muscle is involved in shoulder movement, which is important for detecting techniques such as down-picking, strumming, and chord changes (i.e., left-hand movement when changing chords). We noticed that the anterior right deltoid muscle has slightly higher values, than the lateral, as it tends to be more consistently active during guitar playing; while the *left deltoid* is completely discarded for low activation proved by low values of both ZCR and RMS.

3. **Biceps**: These muscles are involved in elbow movement, which is highly activated during techniques such as bending and vibrato, and they are responsible for elbow flexion strongly activated during string bending and down picking.

As a result, We discarded *the left deltoid and the right brachioradialis*, which fell behind the thresholds defined in 3.2. The hand muscles showed too little SNR due to crosstalk due to higher tissue density concentration in a smaller area.

4.2 Gesture To Sound Mapping

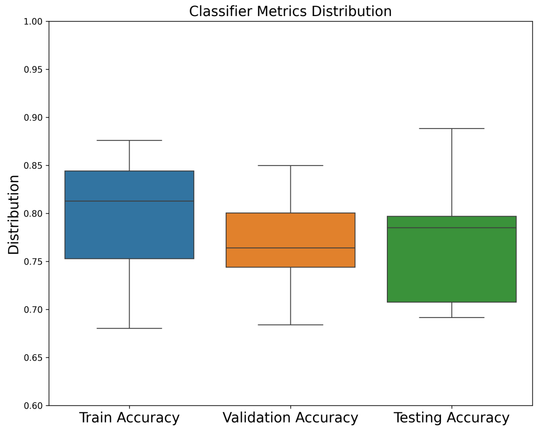
Gesture-to-sound mapping involved two mapping strategies in parallel: one to set the pedalboard preset through gesture classification (between clean, crunch, or heavy) and the other to modulate a set of parameters (e.g., Mix of Reverb, Drive of Overdrive, etc.) by analyzing the level of muscle contraction. As presented in Section 3.5, inside the Max patch we defined three sonic presets: clean, crunchy, and heavy. We associated each with guitar gestures commonly played with that specific sound. The classification model set the preset according to the techniques performed. The regression model modulated six parameters via OSC messages: the overdrive and distortion's drive level, the reverb's mix and room size, and the delay time and chorus' mix.

4.3 Technical Evaluation

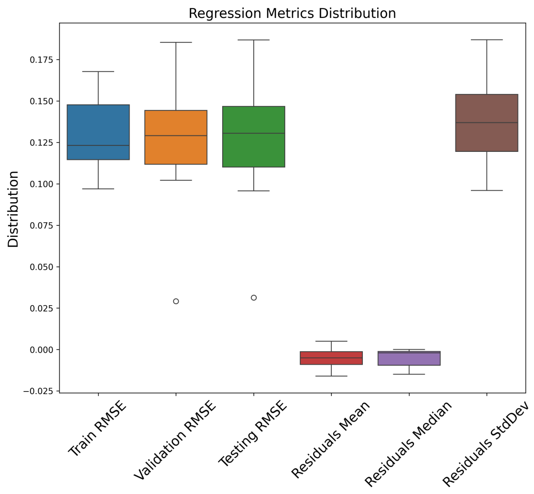
We evaluated the classification and the regression performance using two validation metrics ([15, 5]): accuracy and Root Mean Square Error (RMSE). The system reached a value of *0.9 for accuracy* and *0.09 for RMSE* in the best-performing subject. Fig. 8 presents the two models' metrics' distribution across seven datasets (i.e., the seven participants). The classifier performed a mean validation and testing accuracy of 0.77 (SD = 0.058) and 0.78 (SD = 0.065) respectively, while the regression model performed an average validation and testing RMSE of 0,1378 (SD = 0.029) and 0.1382 (SD = 0.031). There are two positive factors that we can notice from Fig 8: the residuals' means are centered around zero (i.e., they have a normal distribution, hence the regression is working correctly), and the testing accuracy shows a positive skew distribution (i.e., the accuracy values tend to be high).

5. USER STUDY

The experiment involved seven experienced male guitarists (aged between 21 and 37 years old, mean age = 28.25, SD = 5.27, mean height = 177 cm). They had an average musical experience of 15 years. All of them deemed themselves intermediates or experts and no one had previous injury in the upper limbs. All experiments were executed leveraging right-hand solid-body electric guitars, with participants adopting a stationary standing posture, as depicted in Fig. 6. The experiment consisted of the following three stages. To begin, each participant followed the



(a) Classifier's metrics distribution.



(b) Regression model's metrics distribution.

Figure 8: Metrics distribution of the classifier (a) and the regression model (b) across seven datasets.

steps described in Section 3.3.1 to create a training dataset. Subsequently, the participant assessed the system by experimenting with a backing track while alternating between the various techniques employed in the training stage. Finally, a customized questionnaire, designed in compliance with DMI assessment forms developed within the NIME community [4, 27], was used to document the participant's experience. The questionnaire was composed of two parts. The first part consisted of the following nine close-ended questions evaluated on an 11-point Likert scale (coupled with short comments) in order to evaluate some constructs typical of the DMI research community (e.g., playability, expressiveness, effectiveness, etc.):

1. **Enjoyability**: How enjoyable was your experience while trying the system?

2. **Playability**: How much control did you feel while using the tool?

3. **Learnability**: How easy was learning to use it?

4. **Expressiveness**: How much did it help you to enhance your expressiveness?

5. **Novelty**: How much novelty does it introduce to your performance?

6. **Effectiveness**: How much was it able to track your sonic intentions and translate them into an effective combination of effects?

7. **Wearability**: How not invasive do you consider the system?

8. **Adaptability**: How much was the system able to improve following your feedback after the first try?

9. **Usability**: How confident would you feel using it in a live performance?

The second part comprised four open-ended questions:

1. **Main Limitations**: What are the main limitations of the system to use it in a live scenario?

2. **Overall Experience**: How was the experience of using muscular contractions to interact with the pedalboard? Do you think it is a good way to enhance your expressiveness?

3. **Added Value**: What is the added value of the system compared to the standard interaction and how would you integrate it into your artistic practice?

4. **Improvements**: How would you improve the system?

The distribution results of the close-ended questions are shown in Fig. 9. We sorted the key aspects decreasingly by median value to better highlight the items in which the system collected the highest ratings. In the following, we report some of the comments made by participants to motivate their ratings.

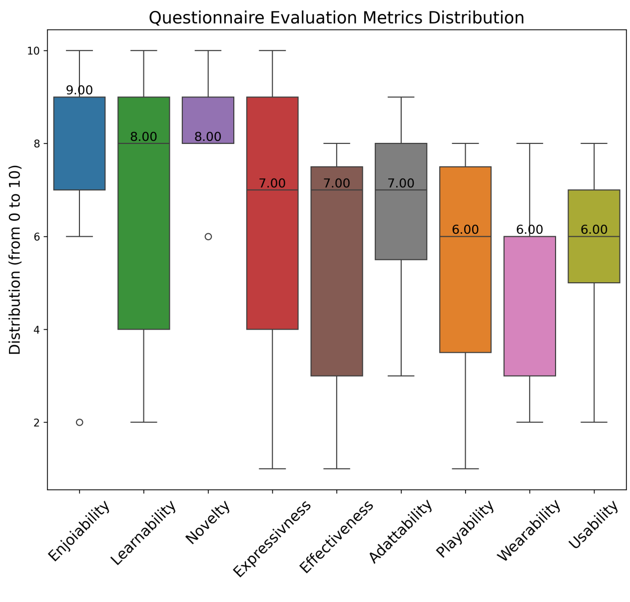


Figure 9: Questionnaire results.

Novelty: all the participants shared a strong appreciation for the idea behind the system, considering it as highly innovative (e.g., "*I believe it would be a great inclusion in improvising, or as a training tool. The system was very interesting*"; "*After an initial moment of general discomfort, I "embraced" the responses of the system and started to play with it, as I was playing along with an (almost) autonomous external agent. This approach introduced a lot of novelties in my playing and I appreciated it.*"; "*Changing sound without iteration with a pedalboard makes the performance more organic*"; "*A musician can concentrate only in the part she* *is playing while exploring the sounds of the pedalboard, interacting with it*"; "*it would make a difference*").

Enjoyability: Five guitarists gave a high grade, while the others a medium one (e.g."*There were times where the system did exactly what I wanted it to do.*"; "*it was cool, I felt super comfortable*") One participant deemed the experience not so enjoyable (e.g. "*I am not very much interested in this kind of technologies. Furthermore, being a first prototype it was not so easy and practical and didn't work so much in my case.*").

Learnability: Four guitarists give a high-grade sharing a fast learning curve (e.g., "*Just as with all guitar instruments and effects it takes a moment to adjust, which in this case is very brief*"; "*Once you try the various techniques and understand the forte-piano dynamics of the device you go like a charm*") one participant complains of too many muscle contractions mapped on small differences in sound ("*The system was intuitive and easy to figure out. However, there is a considerably large number of muscle contractions mapped to different subtle changes in effects. It is difficult to remember which muscle controls which parameter upon the first trial.*)

Expressiveness: Three users gave high grades and two a medium one (e.g., "*The potentials are endless*"; "*In my opinion, it is super creative, if brought to a certain level of safety I would also use it during a live performance.*"). One guitarist who gives a low score said "*the extreme focus of muscle contractions was distracting me from the instrument and my playing*").

Adaptability: It received similar grades to the expressiveness (e.g., "*the more I played it, the more I understood how to use it, the more fun it was*"). One participant gave a valuable feedback to improve the performance by changing the dataset creation procedure: "*My feedback was to train the system letting the player behave as natural as possible, so to avoid having a forced "fake" baseline*". He was against the idea of instructing all the participants to play the same riff for each guitar technique. Another participant complaint about preset change latency, suggesting to collect more data to improve the performance "*What I can say is that the system was overall injecting lots of latency in the process [..] I was able to identify two or three situations in which it follows my techniques [..] It would be interesting to further experiment with it tho, and see if and how the overall accuracy could be improved, possibly recording more data.*"

Playability: Five participants gave a medium or high grade (e.g, "*The fact that effects could have been modulated so easily is mind staggering!*"). Two participants who gave a low grade agreed about a too high muscle exertion needed: "*I had a focus on my muscle contractions a lot to get the system to recognize them. This extreme focus on muscle contractions was a distraction from playing my instrument.*" and the other "*The muscular effort I had to put into the system was quite high: therefore, I lost control over my playing while trying to control the sensors. Indeed, usually, a trained guitar player tends to release muscular tension as much as possible while playing technical passages and concentrate on small, fluent movements.*".

Effectiveness: Two guitarist give an high grade (e.g., "*it behaved really well, I did not expect such incredible efficiency and effectiveness, I had never tried something like this and was impressed*"). Two gave a low grade (e.g., "*I was rarely able to effectively drive the system towards the results I had in mind. Indeed, I preferred to let the system to guide me and play more freely.*")

Usability: One guitarist suggest to change the paradigm by using the system in a different way, not only to change preset and parameters (citing: "*I would probably like to use the system in a quite complex scenario, in a creative and interactive way, rather than switching my pedals [..] a performance entirely based on a system like that, in which the performer can focus towards a complex interaction with the system, would be far more interesting to me.*". Another guitarist appreciated the system for solo concert ("*Very confident in case I play solo, in band contexts it would take more time to adapt*").

Wearability: It is the Achilles heel of the system, two petitioners shared a perception of fragility due to many cables (e.g., "*It was very invasive as you have to wear many sensors.*"); Two others did not have any problem on this side (e.g., "*It did not bother me much wearing the equipment as the experiment proceeded*").

Concerning the open-ended questions, the following considerations were made by participants:

Discomfort. Firstly, all participants mentioned that the main limitation of the system is the poor wearability which creates a sense of insecurity in movement due to the intrusiveness of the sensors and cables and the perceived fragility. Two guitarists mentioned the extreme muscle contraction needed to trigger the classifier.

Experience appreciation. On average, participants considered the overall experience as interesting. However, three of them are skeptical of the classifier due to the latency in changing preset and the extreme contraction needed (e.g., "*It was an interesting experience to use the system. I believe the regression component would be a very interesting tool to use. However, I am skeptical about the classification*". A guitarist reported that the system could be very useful applied to a bilateral interaction not only for the objective of modulating parameters and switching the pedalboard preset (mentioning his words: "*I think that this system could be fruitfully employed in an interactive context, in which the performer dialogues with it, more or less consciously, similarly to what I was trying to do at the end of the experiment. That I think would be particularly interesting*".)

Added Value. Four guitarists appreciated the tracking of the guitar feelings and defined the tool as an innovative way to express creativity and explore art (e.g., "*New ways of interacting with music could lead to new ways to express art and new ways of exploring the art. This aspect can be interesting!*").

Features requests. A major suggestion to improve the system was the enhancement of the normalization strategy to avoid extreme muscle contraction to change presets (e.g., "*I would suggest that you first ask the musicians to play some defined pieces/techniques naturally with the sensors. Thereby you can obtain the minimum, maximum, and most importantly, the median and standard deviations of each muscle contraction. I believe this could eliminate the need for very intense muscle contractions that I experienced during the experiment*").

6. DISCUSSION

Taken together, the results of the user study showed that on average the system was conceived as useful by participants in enhancing their creative process. We would like to highlight a number of technical limitation that affected the system's performance and are helpful for upcoming upgrades.

Unstable classification performance. The unstable classification performance can be attributed to the noisy nature of the sEMG, which leads to a low signal to noise ratio. Crosstalk between adjacent muscles, electromagnetic disturbances, and cable movement artifacts all increase the likelihood of extraneous noise entering the sEMG signal. In this prototype, users suffer from the inconvenience of carrying several not shielded cabled electrodes. LWT3 is developing a new sensor board that incorporates the cables into a comfortable, stretchy, shielded suit in order to address this issue and improve wearability. We experienced a high variance between the best classification accuracy and the worst (0.88 against 0.69). Musicians who initially showed low interest performed the worst, being skeptical about stepping out of their comfort zone during the evaluation phase. Furthermore, the best classification accuracy was achieved when musicians were asked to interpret the given riffs as much as possible (even changing it a bit) to better model their guitar style.

Little training data. The time constraints involved in creating the intra-subject dataset result in little training data, lowering the performance. To achieve the best performance the system has to be trained multiple time by each user independently. For this reason, we plan to add an Interactive Machine Learning (IML) approach to the training process that will allow users to train the system by their own. IML is a supervised learning technique that designs and implements algorithms to ease the learning process with the help of human feedback. One benefit of IML is that the mapping between gesture and sound can be interactively "shown" by the user to the system [13], rather than being manually coded which requires specialized knowledge not often found among musicians [37].

Embedded systems target. We developed the two RNNs to be lightweight (as described in 3.3.3), with the aim of embedding the system into wearable devices with low storage and computational power. Removing this constraint and increasing the models' complexity (i.e., by adding layers) would improve the overall performance.

Latency. The latency exhibited by the system stems from its underlying architectural design. Specifically, the inherent processing mechanism involves two RNNs analyzing a window of 256 samples at a time before generating an inference, resulting in a delay of 256 milliseconds (ms) at a sampling frequency of 1000 Hz. Furthermore, the system requires an additional 50 ms to analyze a new window using standard laptop hardware. Consequently, the cumulative latency for each inference amounts to about 306 ms. Attempts to mitigate latency by reducing the window size proved ineffective as it compromised the system's performance. Theoretically, RNNs benefit from longer windows for enhanced prediction capabilities, thus necessitating a trade-off between predictive accuracy and latency. Through extensive tests, we experimentally determined that a window size of 256 samples yielded the optimal compromise. It is worth noticing that the system is not engineered to achieve real-time, audio-rate latency; rather, its primary objective is to seamlessly adapt the guitar's sound to the musician's gestures. The intention is not to replace standard pedal interactions with faster alternatives, but to introduce an additional method of interacting with effects, capable of dynamically adjusting effect parameters to align with the musician's sonic intentions during performance. The system is intended for use in conjunction with a pedal board, particularly during moments of exploration, improvisation, and soloing. The overarching goal is to facilitate a dialogue between artificial agents and musicians, fostering the emergence of novel effects superposition and thereby enhancing creativity.

Lastly, our study also presents a limitation of the relatively low participants number. A larger pool of participants could have potentially provided other insights. A larger pool of participants, would allow one to overcome the intra-user nature of the system, and train it in an inter-user manner with a single big dataset. Furthermore, our study involved only males. A larger number of participants and a more gender-balanced study would make our results more generalizable. Despite these limitations, the achieved results suggest that the system was effective in demonstrating the potential of the proposed interaction strategy. Our plan is to extend this protocol to other instruments and musicians in the future.

7. CONCLUSIONS

The objective of this research was to introduce a system that provides guitar players with a new way to interact with their sound effects. Specifically, we aimed to investigate how muscle contractions can be useful to guitar players as a new form of expression with respect to the conventional use of effects pedalboards. To enhance the standard pedal interaction, we developed a DMI that automatically adjusts the sound according to the performed gesture and exerted force. With the proposed system guitarists can modulate different effects' parameters by changing the amount of muscular contraction. The muscle contractions are thus used as an expression pedal. This approach is in contrast with the typical guitar pedalboard usage, where the guitarist must set all the effects parameters prior to a performance.

We presented a selection of the best muscle groups for guitar gesture classification, to increase the classification performance. By leveraging this selection, the system was able to control a set of sound effects according to the user's sound intention (after an initial training stage), consequently changing the sound of the guitar during a performance.

To evaluate the system, we conducted a user study with seven experienced guitarists. Overall, results suggested that the proposed form of interaction can represent an alternative means to change the sound of the guitar, providing guitar players with a creative way to enhance their expressiveness.

In future work, we plan to extend the protocol to other musicians (e.g pianist or wind instruments) leveraging the potential of full body mapping offered by the custom acquisition board developed by our partner LWT3. Furthermore, we plan to increase the classification's robustness, by utilizing the sEMG in conjunction with inertial measurement unit's sensors.

Leveraging the Internet of Musical Things paradigm [35], we also plan to extend the use of our method to the real-time control of different types of musical devices wirelessly connected to the sEMG wearable device (e.g., smoke machines, stage lights), thus giving musicians an additional means to express themselves and create novel forms of live music performances.

8. ETHICAL STANDARDS

All the involved participants provided informed consent. There are no observed conflicts of interest in this study.

9. REFERENCES

[1] P. K. Artemiadis and K. J. Kyriakopoulos. Emg-based control of a robot arm using low-dimensional embeddings. *IEEE transactions on robotics*, 26(2):393–398, 2010.

[2] R. B. Azhiri, M. Esmaeili, and M. Nourani. Real-time emg signal classification via recurrent neural networks. In *2021 IEEE International Conference on*

Bioinformatics and Biomedicine (BIBM), pages 2628–2635. IEEE, 2021.

[3] M. Barbero, R. Merletti, and A. Rainoldi. *Atlas of Muscle Innervation Zones: Understanding Surface Electromyography and Its Applications*. Springer Milan, 2012.

[4] J. Barbosa, J. Malloch, M. M. Wanderley, and S. Huot. What does" evaluation" mean for the nime community? 2015.

[5] G. Bellocchi, M. Rivington, M. Donatelli, and K. Matthews. Validation of biophysical models: issues and methodologies. a review. *Agronomy for Sustainable Development*, 30(1):109–130, 2010.

[6] D. Buongiorno, G. D. Cascarano, A. Brunetti, I. De Feudis, and V. Bevilacqua. A survey on deep learning in electromyographic signal analysis. In *Intelligent Computing Methodologies: 15th International Conference, ICIC 2019, Nanchang, China, August 3–6, 2019, Proceedings, Part III 15*, pages 751–761. Springer, 2019.

[7] F. Calegario, M. M. Wanderley, S. Huot, G. Cabral, and G. Ramalho. A method and toolkit for digital musical instruments: Generating ideas and prototypes. *IEEE MultiMedia*, 24(1):63–71, 2017.

[8] B. Di Donato, J. Bullock, and A. Tanaka. Myo mapper: a myo armband to osc mapper. 2018.

[9] B. Di Donato, A. Tanaka, M. Zbyszynski, and M. Klang. Eavi emg board. 2019.

[10] L. Donovan and A. P. McPherson. The talking guitar: Headstock tracking and mapping strategies. In *NIME*, pages 351–354, 2014.

[11] C. Erdem, Q. Lan, J. Fuhrer, C. P. Martin, J. Tørresen, and A. R. Jensenius. Towards playing in the'air': Modeling motion-sound energy relationships in electric guitar performance using deep neural networks. In *Proceedings of the SMC Conferences*, pages 177–184. Axea sas/SMC Network, 2020.

[12] C. Erdem, B. Wallace, and A. R. Jensenius. Cavi: A coadaptive audiovisual instrument–composition. In *NIME 2022*. PubPub, 2022.

[13] J. Françoise, N. Schnell, and F. Bevilacqua. Mad: mapping by demonstration for continuous sonification. In *ACM SIGGRAPH 2014 Studio*, pages 1–1. 2014.

[14] R. I. Godøy and M. Leman. *Musical gestures: Sound, movement, and meaning*. Routledge, 2010.

[15] M. Grandini, E. Bagli, and G. Visani. Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*, 2020.

[16] H. Hewamalage, C. Bergmeir, and K. Bandara. Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1):388–427, 2021.

[17] I. K. M. Jais, A. R. Ismail, and S. Q. Nisa. Adam optimization algorithm for wide and deep neural network. *Knowledge Engineering and Data Science*, 2(1):41–46, 2019.

[18] M. Janke, M. Wand, and T. Schultz. A spectral mapping method for emg-based recognition of silent speech. In *B-Interface*, pages 22–31, 2010.

[19] A. R. Jensenius, V. E. Gonzalez Sanchez, A. Zelechowska, and K. A. V. Bjerkestrand. Exploring the myo controller for sonic microinteraction. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 442–445. Aalborg University Copenhagen, 2017.

[20] C. Jorgensen, D. D. Lee, and S. Agabont. Sub auditory speech recognition based on emg signals. In *Proceedings of the International Joint Conference on Neural Networks, 2003.*, volume 4, pages 3128–3133. IEEE, 2003.

[21] K. Kiatpanichagij and N. Afzulpurkar. Use of supervised discretization with pca in wavelet packet transformation-based surface electromyogram classification. *Biomedical Signal Processing and Control*, 4(2):127–138, 2009.

[22] M. Leman. *Embodied music cognition and mediation technology*. MIT press, 2007.

[23] T. Matsubara and J. Morimoto. Bilinear modeling of emg signals to extract user-independent features for multiuser myoelectric interface. *IEEE Transactions on Biomedical Engineering*, 60(8):2205–2213, 2013.

[24] E. A. Meneses, S. Freire, and M. M. Wanderley. Guitarami and guiart: two independent yet complementary augmented nylon guitar projects. In *Proceedings of NIME*, 2018.

[25] K. Nymoen, M. R. Haugen, and A. R. Jensenius. Mumyo–evaluating and exploring the myo armband for musical interaction. 2015.

[26] B. Potočnik, M. Divjak, F. Urh, A. Frančič, J. Kranjec, M. Šavc, I. Cikajlo, Z. Matjačić, M. Zadravec, and A. Holobar. Estimation of muscle co-activations in wrist rehabilitation after stroke is sensitive to motor unit distribution and action potential shapes. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(5):1208–1215, 2020.

[27] P. J. C. Reimer and M. M. Wanderley. Embracing Less Common Evaluation Strategies for Studying User Experience in NIME. In *NIME 2021*, apr 29 2021. https://nime.pubpub.org/pub/fidgs435.

[28] C. Rhodes, R. Allmendinger, and R. Climent. New interfaces and approaches to machine learning when classifying gestures within music. *Entropy*, 22(12):1384, 2020.

[29] J. A. Sandoval-Espino, A. Zamudio-Lara, J. A. Marbán-Salgado, J. J. Escobedo-Alatorre, O. Palillero-Sandoval, and J. G. Velásquez-Aguilar. Selection of the best set of features for semg-based hand gesture recognition applying a cnn architecture. *Sensors*, 22(13):4972, 2022.

[30] A. Tanaka. Intention, effort, and restraint: The emg in musical performance. *Leonardo*, 48(3):298–299, 2015.

[31] A. Tanaka, B. Di Donato, M. Zbyszynski, and G. Roks. Designing gestures for continuous sonic interaction. 2019.

[32] A. Tanaka and M. Donnarumma. The body as musical instrument. *The Oxford handbook of music and the body*, pages 79–96, 2019.

[33] A. Tanaka, F. Visi, B. Di Donato, M. Klang, and M. Zbyszyński. An end-to-end musical instrument system that translates electromyogram biosignals to synthesized sound. *Computer Music Journal*, pages 1–40, 2024.

[34] A. Toro-Ossaba, J. Jaramillo-Tigreros, J. C. Tejada, A. Peña, A. López-González, and R. A. Castanho. Lstm recurrent neural network for hand gesture recognition using emg signals. *Applied Sciences*, 12(19):9700, 2022.

[35] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet. Internet of Musical Things: Vision and Challenges. *IEEE Access*, 6:61994–62017, 2018.

[36] F. Visi, B. Caramiaux, and M. Mcloughlin. A Knowledge-based, Data-driven Method for Action-sound Mapping. In *NIME 2017 : New Interfaces for Musical Expression*, Copenhagen, Denmark, May 2017.

[37] F. G. Visi and A. Tanaka. Interactive machine learning of musical gesture. *Handbook of artificial intelligence for music: Foundations, advanced approaches, and developments for creativity*, pages 771–798, 2021.

[38] M. Wright. Open sound control: an enabling technology for musical networking. *Organised Sound*, 10(3):193–200, 2005.

[39] M. Zbyszyński, B. Di Donato, F. G. Visi, and A. Tanaka. Gesture-timbre space: Multidimensional feature mapping using machine learning and concatenative synthesis. In *Perception, Representations, Image, Sound, Music: 14th International Symposium, CMMR 2019, Marseille, France, October 14–18, 2019, Revised Selected Papers 14*, pages 600–622. Springer, 2021.