Fatigue detection in EMG signals.

Inês Freitas
Advisor: Prof. Ana Fred

Index Terms—Electromyography, muscular fatigue, pattern recognition, signal processing

I. INTRODUCTION

Fatigue in athletes is, sometimes the cause of injuries. The detection of fatigue is done when the athlete is already injured. In this thesis, the goal is to detect fatigue in athletes based on the analysis of EMG (Electromyography) signals. The EMG signals, resulting from impulse-like electrical muscle activity, are acquired through surface electrodes. In this work we focus on the rowing activity. It was chosen because it is a very complete activity and the results obtained with it can be extended to other activities. The signals used in this work were provided by the “Faculdade de Motricidade Humana”, where the acquisition of the signals took place. In this case each athlete performed 3 tasks, using the Concept2, an indoor rower that simulates the resistance of the water. The principal task is a 2000 meters proof that normally takes approximately 6-8 minutes [1]. In this task athletes do their best at a maximum effort. Before and after the main task the rowers had done 10 cycles at maximum effort. The three tasks performed by the individuals are mentioned as PRE2000 (10 cycles before the 2000 meters task), 2000 (the 2000 meters task) and POS2000 (10 cycles after the 2000 meters task). To acquire the signal, surface electrodes were used in the main muscles involved in the action of rowing. The muscles selected to this study are Posterior Deltoid (PD), Vastus Lateralis (VL), Biceps Femoris (BF) and Biceps Brachii (BB). In this thesis we propose an analysis of the dynamic of the pattern of the power evolution across time/active zone, corresponding to a cycle of the EMG signal. Additionally it was also performed an analysis of the force signal.

II. OVERVIEW OF EMG ANALYSIS

The electromyographic signal (EMG) is a electrophysiological signal that measures electrical currents generated in muscles during its contraction [2]. Invasive and non-invasive methods have been used to acquire the electrical impulses generated by muscles. In the invasive method a needle is inserted directly into the muscle through the skin. The non-invasive way is recorded with electrodes on the skin surface. The surface EMG signal is achieving more importance in several fields, being used for the diagnosis of muscle or nerve disorders, and for the analysis of the neuromuscular system [3]. Studying EMG signals can also help controlling prosthesis. The EMG signal can also be used to detect muscle fatigue. It is very important to detect this condition because early detection can avoid irreversible injuries to the subject [4]. The muscle fatigue is considered the incapacity to maintain the desirable level of force performing a specific task. Using the EMG signal, to detect muscle fatigue, the parameters normally used are the amplitude and the frequency of the signal. The studies in the literature show that the amplitude of EMG signals increases progressively as a function of time when the fatigue increases. Other studies show that the mean power frequency decreases as a function of time [5].

According to [5] muscle fatigue can be monitored by changes in the EMG frequency properties such as mean and median frequency. In this thesis, we propose to analyze the EMG signal by using the signal divided into cycles. The analysis proposed is based on the evolution of the mean power across each cycle, which is used to detect fatigue and to identify individuals. Fatigue is detected by the classification of each cycle according to the task involved and the individuals are identified by classifying each cycle according to the individuals.

III. PROPOSED METHODOLOGY

The diagram in figure 1 gives a schematic description on the steps of this work. First it is done the denoising of the signal in order to eliminate anomalous spikes. Then both EMG and force signals are divided into segments, for the force signal those segments are called cycles while for the EMG signal are called active zones. The cycles of the force signal are
used to help dividing the EMG signal into active zones. Then it is computed the spectrogram of each active zone from which power and frequency related features are computed, such as mean power evolution of each active zone. Based on the mean power evolution it is performed the task identification (to analyze fatigue in individuals) and the human identification (to identify the individuals according to the mean power evolution of each active zone). The mean power is used because it has a typical temporal pattern for all the individuals.

Figura 1. Schematic description of the steps of this work.

In figure 2 can be seen an example of the EMG and force signal, as well as an active zone and a cycle of the force signal. A cycle can be defined as a set of movements which are repeated during the activity. Active zones correspond to the period in which the muscle is active. Each active zone from the EMG signal correspond to a part of the corresponding cycle in the force signal.

In this work it is expected to analyze each active zone separately and then compare the results achieved. The base of all the analysis was done with the spectrogram of each active zone. Let $P(t_i, f_j)$ be the power spectrum density at instant $t_i$ and frequency $f_j$ for a given active zone. Based on the spectrogram, two kinds of features were calculated: global features characterizing the global behavior of each active zone; and features associated with the temporal evolution along each active zone. The techniques used to calculated those features are explained next.

The analysis made in this work is based on the active zones of each muscle. The observation of the active zones allows us to see that the first active zone sometimes is incomplete as well as the last one. So the work done excludes these two ones to avoid compromising the results.

A. Global Features

Global features were calculated for each active zone. These features are important to see the evolution of fatigue along each task. The most important parameters to analyze are the power and the frequency. It was calculated the maximum and the mean power of each active zone and the dominant frequency of the active zone. The dominant frequency, $F_{dom}$, is the frequency associated with the maximum power value, as shown in equation 1.

$$F_{dom} = \arg_{t_i, f_j} \max_{t_i, f_j} P(t_i, f_j)$$

The mean power of each active zone, $P_{mean_{global}}$, is calculated according to equation 2.

$$P_{mean_{global}} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} P(t_i, f_j)$$

In the equation (2), $n$ is the number of samples considered in the calculation.

B. Dynamic temporal evolution

In order to see the temporal evolution along each active zone it is important to analyze each instant of time of each active zone. So we calculated features associated with each instant of time $t_i$. The parameters calculated were the dominant frequency (frequency associated with the maximum power) $F_{dom_i}$, the maximum power $P_{max_i}$, the mean power $P_{mean_i}$, the mode power (value that occurs more frequently), the median power (number which separates the higher half of a sample) and the standard deviation of the power.

$$F_{dom_i} = \arg_{f_j} \max_{f_j} P(t_i, f_j)$$

$$P_{max_i} = \max_{f_j} P(t_i, f_j)$$

$$P_{mean_i} = \frac{1}{n} \sum_{j=1}^{n} P(t_i, f_j)$$

Figura 2. Example of an EMG and force signals, as well as the active zone and cycle respectively, associated with a rowing activity.
The evolution of the mean power along each active zone presents a typical pattern, as illustrated in figure 3. It consists of an initiated period with high mean power values followed by a period where the mean power decreases abruptly, and an ending period where the mean power stays in lower values. For all the individuals it can be seen the same pattern, where changes from individual to individual and task to task is the instant when the mean power decreases or the duration of the active zone. For those reasons, this was the pattern chosen to continue the analysis. This result is more relevant in the first channel, where the graphics of the evolution of this feature have a more consistent pattern. According to this fact the proposed work has focused on the muscle named Posterior Deltoid.

![Graph of mean power evolutions](image1)

**Figure 3.** Evolution of the mean power along the time, $P_{mean}(t_i)$, in three active zones, for individual 1, channel 1.

Based on all the information collected and the results obtained one can say that each individual could have a different pattern of active zone and also each task could also have some different characteristics from the others tasks. According to this purpose, analyzing each active zone, one tried to identify the individual and the task of each active zone. The algorithm used was the k-Nearest Neighbor (7-NN), which classifies signals based on the Euclidean-distance between feature vectors.

This algorithm was used to classify the signals in two ways. First to see to which task correspond the signal and then to see to which individual correspond the signal. The methodologies used for the two classifications are presented next.

The features used in the task and individual classification were the instant of change ($t_{ch}$), the duration of the active zone ($d_{act}$), and the area under the curve. The value of $t_{ch}$ was considered as the minimum of the second derivative of the active zone in region of the maximum of the signal, and was calculated for all the active zones concerned to the tasks PRE2000 and POS2000, for channel 1. It was also calculated the mean value of the change instant, $\bar{t}_{ch}$=0.392, as well as the standard deviation of the same instant, $\sigma_{t_{ch}}$=0.0967, using both PRE2000 and POS2000 signals of all individuals. The area under the curve is calculated until different values as shown next by the different scenarios.

1) **Task classification - Fatigue related analysis:** As it was mentioned before, tasks PRE2000 and POS2000 are similar to each other. They both consist in performing 10 rowing cycles at the maximum effort. Each athlete has his own way of doing those tasks but it is expected that they perform the two tasks with the same technique. However fatigue could change the EMG signal and, consequently, the active zones from task PRE2000 could be different from those achieved in task POS2000. The algorithm k-NN was used in 4 different scenarios, named as A, B, C and D. Those scenarios and corresponding features used are explained next. The difference between the scenarios is in the way the area is calculated.

a) **Scenario A:** In the first scenario were used three features: the area under the curve ($A_A$), the change instant ($t_{ch}$) and the duration of the active zone ($d_{act}$). The $A_A$ was calculated until the end of the first part of the active zone. To define this instant two experiments were made, called A1 and A2.

Experiment A1 was calculating the area $A_A$ until $t_{ch}$, this was the first feature used. Next the same algorithm was used but with two features, the area mentioned before and $t_{ch}$ of each active zone. At last another feature was added which was $d_{act}$ and the algorithm was used with the three features. In figure 4 can be seen an example of these features, are identified the $t_{ch}$ for each active zone and the line in black shows the value of $t_{ch}$.

![Area under curve examples](image2)

**Figure 4.** Example of the area $A_A$ for task POS2000 (green).

Experiment A2 was using the area until the mean value of the change instant minus the standard deviation ($t_{ch}$-$\sigma_{t_{ch}}$), this was the first feature used. The features used next were the same as in experiment A1, and can be seen an example in figure 4, the only difference is that the line in black takes a lower value, but it is not visually significant.

b) **Scenario B:** In scenario B were used two features: the area under the curve ($A_B$) and $t_{ch}$. The difference from the scenario A is that the area is calculated differently. In this case to choose until where the area was calculated it was compared the two correspondent active zones of each task, for example, the second active zone of task PRE2000 was compared with the second active zone of task POS2000. Comparing $t_{ch}$ of the two active zones it was chosen the maximum of the two. The area was calculated until this value minus the $\sigma_{t_{ch}}$, for the two active zones correspondent to the two tasks. In order to calculate this area both time and power were normalized by the maximum, as can be seen in figure 5.

In figure 5 we see that the curve in blue has a higher $t_{ch}$ so was that one that was used. First the algorithm was used with just one feature which was $A_B$. Then it was used the two features together, $A_B$ and $t_{ch}$.

c) **Scenario C:** In scenario C the feature used was the area under the curve ($A_C$), calculated until the $t_{ch}$ of each
active zone minus $\sigma_{t,ch}$. In order to calculate this area both time and power were normalized by the maximum as shown in figure 6.

![Figure 5](image1.png)
**Figura 5.** Example of the area $A_B$ for the task POS2000 (green).

![Figure 6](image2.png)
**Figura 6.** Example of the area $A_C$ for the task POS2000 (green).

**d) Scenario D:** In scenario D the method used was slightly different. It was calculated the correlation index of all the signals, of the same task, for each individual, taking as reference the second active zone. Based on the correlation indexes was calculated the shift, of each active zone, when compared to the second active zone. With the deviation mentioned before the active zones were shifted according to the value of the deviation. After this procedure the active zones were aligned as can be seen an example in figure 7.

![Figure 7](image3.png)
**Figura 7.** Example of the alignment of the active zones for one individual performing one task.

After applying this method all the active zones were aligned, then it was calculated the mean active zone and then $t_{ch}$ of it. The goal of this method was to have the instant to calculate the area under each curve. The area ($A_D$) was calculated, for each active zone (without the alignment), until the $t_{ch}$ mentioned before. For this process both time and power were normalized. In scenario D was used just one feature: $A_D$.

2) **Individual Identification:** As it was mentioned before, each individual has his own way of performing the tasks, so it could be possible to identify an individual by his active zones. The active zones acquired in the process were classified using the k-NN algorithm to identify the individuals. Two features were used, which are the duration of the active zone and the change instant ($t_{ch}$).

The algorithm was used with each feature separately and with both simultaneously. This means that 3 scenarios were used, one with $d_{act}$, other with the fall instant and other with these two features together. For the 3 scenarios was applied the algorithm using just the signals from the 3 tasks isolated and then the signals from the 3 tasks simultaneously.

IV. RESULTS AND DISCUSSION

The population consists of 11 athletes that have been doing rowing for several years, mentioned in this work by numbers from 1 to 11.

To acquire the EMG signal were used surface electrodes in the main muscles involved in the task of rowing, as mentioned in table I. It was also obtained the force signal for each task and each individual.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Muscles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Posterior Deltoid</td>
</tr>
<tr>
<td>2</td>
<td>Vastus Lateralis</td>
</tr>
<tr>
<td>3</td>
<td>Biceps Femoris</td>
</tr>
<tr>
<td>4</td>
<td>Biceps Brachii</td>
</tr>
</tbody>
</table>

To identify the muscle that is being in focus it is mentioned the channel where it was recorded.

A. Fatigue Analysis

1) **Global features:** As mentioned in chapter 2 the literature say that the amplitude of EMG signals increases progressively as a function of time when the fatigue increases and the mean power frequency decreases as a function of time [5]. In this study is proposed another method to identify fatigue as explained in chapter 3.

The global features, detailed in section 3.4.1, are calculated for each active zone and then compared to see the evolution of the feature along each task. These provide a global overview of the evolution of the variation of the power and frequency during each task. The results obtained show that there is not a constant behavior with these features.

In figure 8 can be seen the mean power of each task for each individual. The blue line corresponds to the values of task PRE2000 for each individual, while the red line corresponds to task POS2000. As it can be seen task PRE2000 has higher values for the mean power. As the two tasks are similar the only difference between them is the effort made between them, consequently the fatigue. If it was used a classifier based on $\Delta Pmean > 0$, where $\Delta Pmean$ is the difference between $Pmean$ of task PRE2000 and $Pmean$ of task POS2000, we would obtain an error probability of 13.6%.
2) Dynamic temporal evolution: Time related features are considered the features related to each instant of time inside each active zone. These ones are useful to see how the parameters evolve across the active zone. The more relevant results were achieved with this analysis.

All the graphics achieved for all the individuals, of the mean power evolution, exhibit the same pattern. It consists of a period of the active zone where the mean power has higher values, followed by a period where the mean power decreases, and at the end a period where the mean power stays in lower values. The pattern is the same but from individual to individual and task to task some parameters change. For example a parameter that changes from individual to individual is the duration of the active zone. Some athletes have longer active zones than others. Another parameter that differs is the fall instant, considered the instant when the mean power decreases, which tends to be similar for the same individual but different when comparing two or more individuals. These two instants are very important to the work performed and they were used in the identification of the individuals and tasks.

Some examples can be mentioned according to the difference between individuals. For example the individual 2 has longer active zones than individual 1. Individual 9 is the one that the mean power stays less time in high values. Individual 6 seems to have the higher values of mean power for the first part of the active zone. All of these evidences show the individuality of each athlete. Based on this fact it was tried to identify individuals by their active zones, the results achieved are presented in the next section.

B. Identifying Individuals and Tasks

1) Tasks classification: The results achieved in task classification are summarized in confusion matrices. It was calculated confusion matrices for each individual and a global one which include all the individuals. The rows correspond to the real class of the signal, in this case the task that the signals correspond to. The columns refer to the task to which the signal was classified.

\[a) \text{Scenario A:} \] In the first scenario were used three features: the area $A_A$, $t_{ch}$ and $d_{act}$, in two experiments.

\[\text{Experiment A1} \]

Experiment A1, as mentioned before, was calculating the area until $t_{ch}$, this was the first feature used. Next was used two features, the area mentioned before and $t_{ch}$ of each active zone. At last was added $d_{act}$.

The global results achieved with the first feature, the area, are presented in the table II.

<table>
<thead>
<tr>
<th></th>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE</td>
<td>57,1429</td>
<td>42,8571</td>
</tr>
<tr>
<td>POS</td>
<td>45,4545</td>
<td>54,5455</td>
</tr>
</tbody>
</table>

Using just the area as feature, and analyzing the results, it is observed that the results are different according to each individual. Commonly a signal is characterized by different parameters and it is the conjunction of several of them that represents the signal. The results presented in table II composed averaging of all the individuals. It can be seen that the percentage of a correct identification is about 55% which means that more than 50% of the tasks are well identified as PRE2000 or POS2000. This means that the area contains discriminating power about the active zones, but it should be used in conjunction with other ones in order to obtain better performance.

The global results for all the individuals, using as features the area and the $t_{ch}$, are presented in table III.

<table>
<thead>
<tr>
<th></th>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE</td>
<td>74,0260</td>
<td>25,9740</td>
</tr>
<tr>
<td>POS</td>
<td>29,8701</td>
<td>70,1299</td>
</tr>
</tbody>
</table>

In this case the results, in general, are better than using just the area as feature. Comparing the global matrix, in table III, with the one in table II, one concludes that the results are improved. The percentage of a correct task identification goes from approximately 55% to above 70%. This means that the two features are important to characterize the active zones, and differentiate them according to the tasks.

Finally, another feature was added which was $d_{act}$. In table IV are presented the average results for all the individuals, using the three features.

<table>
<thead>
<tr>
<th></th>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE</td>
<td>94,8052</td>
<td>5,1948</td>
</tr>
<tr>
<td>POS</td>
<td>5,1948</td>
<td>94,8052</td>
</tr>
</tbody>
</table>
Analyzing the results achieved, using one, two and three features, one sees that, with the use of the three features the results are better. In this case the majority of the tasks are well classified, one have approximately 95% of tasks well classified. The three features together is the best way to have a good identification of the tasks of the active zones in this experiment.

**Experiment A2**

Experiment A2, as referred before, was using the area until $t_{ch}$ minus $\sigma t_{ch}$, this was the first feature used. The features used next were the same as in experiment A1.

The global results, using the area as feature, for each individual, are shown next in table V.

<table>
<thead>
<tr>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>55,8442</td>
<td>44,1558</td>
</tr>
<tr>
<td>44,1558</td>
<td>55,8442</td>
</tr>
</tbody>
</table>

Comparing experiment A1 and A2, using just one feature, the results are not conclusive in which is the best choice. With experiment A1 the results are better for some individuals and with experiment A2 are better for other individuals. Comparing the global matrices, the results are very similar and it is not possible to get a good conclusion, but as it was mentioned before one feature, usually is not enough to characterize a signal.

In tables VI and VII are presented the global results, for 2 features and 3 features, respectively.

<table>
<thead>
<tr>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>75,3247</td>
<td>24,6753</td>
</tr>
<tr>
<td>23,3766</td>
<td>76,6234</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>94,8052</td>
<td>5,1948</td>
</tr>
<tr>
<td>3,8961</td>
<td>96,1039</td>
</tr>
</tbody>
</table>

As it was expected the results are improved by the use of more than one feature. It is interesting to see that with this experiment the global results with two and three features are sightly better in experiment A2. But it can not be considered a significant difference so it can not be said that on experiment is better than the other. In some individuals is best with one experiment and for others with the other experiment.

**b) Scenario B:** In scenario B were used two features: the area and the change time instant. The global results are presented in table VIII.

<table>
<thead>
<tr>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>58,4416</td>
<td>41,5584</td>
</tr>
<tr>
<td>41,5584</td>
<td>58,4416</td>
</tr>
</tbody>
</table>

The global matrix (see figure VIII) shows that all the individuals together have approximately 58% of a correct task classification.

With the use of two features the results, in general, get better. The percentages of active zones correctly classified are higher than with just one feature. This fact is also observed in the global matrix, figure IX, where the percentage of a correct classification is approximately 72%, showing that those two feature are indicated to identify the tasks.

**c) Scenario C:** In scenario C was used one feature: the area under the curve, as explained in section 3.4.2.

The global results of scenario C with just one feature (area) are presented in table X.

<table>
<thead>
<tr>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>61,0390</td>
<td>38,9610</td>
</tr>
<tr>
<td>38,9610</td>
<td>61,0390</td>
</tr>
</tbody>
</table>

Analyzing the table one see that it were achieved good results just with one feature. The global matrix shows that 61% of the active zones are correctly classified, which comparing with the use of just one feature in the other scenarios, is a good result. In this case it was not possible to use more than one feature but as mentioned before the better results, with this kind of algorithm, are achieved with more than one feature.

**d) Scenario D:** In scenario D was used just one feature: the area. The global results are presented next in table XI.

<table>
<thead>
<tr>
<th>PRE</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>79,2208</td>
<td>20,7792</td>
</tr>
<tr>
<td>27,2727</td>
<td>72,7273</td>
</tr>
</tbody>
</table>

With just one feature, one see that the results are very good, compared with the other scenarios. The global matrix shows
that active zones have about 75% of being identified with the correct task. This could mean that the active zones of the same task could have a similar pattern but shifted.

e) **Comparison of the 4 scenarios:** All the 4 scenarios were attempts to reach the best results so it is important to compare them. As it was mentioned before the percentages achieved with just one feature are not so good as using more features. Comparing the use of just one feature in the 4 scenarios, scenario D is the best one because the percentage of a good classification is very high. Even compared with the other scenarios but using two features the results of scenario D are similar. So it can be said that scenario D is the best one. This result let us to conclude that, probably, the different active zones from the same task are shifted as time evolves. A reason for that shift could be fatigue because the pattern of the individual is the same but due to fatigue the active zones are not equal to each other.

With the two features the results are improved, in some cases significatively but it is with the three features that one has the best results and almost all the active zones are correctly classified. This fact proves that all the features are relevant for this study and the combination of the three identifies the task in focus.

Another result that can be discussed is the fact that active zones from PRE2000 and POS2000 are different. Theoretically PRE2000 and POS2000 are two similar tasks in which athletes should produce the same results. But in fact the results are different, probably due to fatigue. Although the tasks are equal, the athletes are not in the same conditions. When they perform POS2000 they had already done another task, the 2000, so fatigue can explain the difference in the active zones from task PRE2000 from POS2000.

2) **Individuals classification:** The active zones acquired in the process were classified using the k-NN algorithm to identify the individuals. Several scenarios were considered and two features were used, which are the duration of the active zone and the time of fall, both mentioned before.

The algorithm was used with each feature separately and with both simultaneously. This means that 3 scenarios were used, one with the duration of the active zones, other with the time of fall and another with these two features together. For the 3 scenarios was applied the algorithm using just the signals from the 3 tasks isolated and then the signals from the 3 tasks simultaneously.

The first results are related to the duration of the active zone. Comparing the results related to the separated tasks one can see that the task which leads to the worst results is the 2000. This can be explained because this task is longer than the others so the active zones could have different patterns inside this task. To better explain one can say that, for example, at the beginning of the task individuals can have active zones with a different pattern from the active zones at end of the same task. Using the tasks PRE2000 and POS2000 individually one achieved better results. One see that, in those cases, it is possible to identify several individuals by their active zones. For example using the active zones from task PRE2000 the individual 4 is completely identified by his active zones and the individual 1 and 2 also have a good percentage of identification. Using the task POS2000 it is the individual 8 which has a 100% correct identification, although individuals 1, 2 and 6 also have a good percentage. When it is joined the three tasks, as most of the active zones are from task 2000, the results are very similar to the ones achieved with this task.

Results concerned to the time of fall show that using task 2000 one obtains the same behavior, explained before, for the other feature. Concerning the other two tasks one see again that some individuals are well classified. Using the active zones from task PRE2000 one sees that individual 6 is better identified than the others, while using task POS2000 it is individual 4 that achieves best results. The joining of the three tasks lead us to the same conclusions presented in the previous paragraph.

Comparing the results achieved with the two features, one sees that the individuals are better classified by the duration of the strokes. So this feature is more related to the individuality of each signal than the other feature. It is expected however that, when the 2 features are considered, one obtains better results.

The results achieved using the 2 features simultaneously we have better results and more individuals are correctly classified. And also considering the task 2000 the results are improved. Using the active zones from task PRE2000 one sees that 5 individuals have percentages of a correct identification higher than 50%. But it can be seen that the individuals are better identified by the active zones from task POS2000. In this one 7 individuals have percentages higher or equal to 50%. With task 2000, in this case it can be seen an improvement because, comparing with the use of just one feature, in this case one has a higher number of active zones correctly identified. Consequently using all the tasks together it is possible to see that the results are better than using just one feature. Although one has higher percentages when the identification is done with active zones taken from tasks PRE2000 and POS2000.

V. **Conclusion**

The goal of this thesis was to analyze the EMG signal of athletes performing rowing. The purpose of the analysis was to identify and characterize fatigue. It was also studied, trough the use of pattern recognition techniques, the individuality of the signals. Another approach was the classification of the active zones. Using pattern recognition techniques, the active zones were classified according to the individual who performed it and according to the task. Using the k-NN algorithm was made the classification of the active zones by its task, 4 scenarios were considered. The difference between the scenarios was the way the area, above the curve of the mean power, was calculated. The results achieved allow us to conclude that tasks PRE2000 and POS2000 are different from each other, even though the nature of the two tasks is similar. Fatigue could be an explanation for this fact, because task POS2000 is performed when the athlete had already made a significant effort. The active zones were also used to identify individuals, using the k-NN algorithm. The results achieved let us conclude that each individual is unique and has a unique
way to perform the activity. So it is possible to identify the individuals by its active zones, using the right features. For future work, one would suggest the study of the other channels (muscles). Since all the work was done for the first channel it is also important to see the behavior of active zones for the other muscles. The techniques should be slightly different because the signals achieved in the other channels have a different behavior. It is also suggested a more exhaustive study of task 2000. One suggests the division of task 2000 in smaller signals to easily analyze the signal and study the different parts of it. Another approach could be the use of other features to improve the classification of tasks and individuals.

**Referências**