Sleep Posture Monitoring

Giovanni Gazzola
School of Internet and Multimedia
University of Padua

Nicola Gastaldello
School of Internet and Multimedia
University of Padua

***Abstract*—Sleep posture analysis is widely used for clinical patient monitoring and sleep studies. The correct and automatic detection of a patient and the way he sleeps can be very usefull to take care of him. In this study, the classification is done using data acquired by a commercial pressure map system from the PmatData dataset. This given data is preprocessed in order to reduce its dimension. Not only the classification accuracy of the model is taken into consideration but also the total amount of data that is needed to perform it. Two way of building a new more compact dataset are presented, the first one using just a Principal-Component-Analysis (PCA) over the data and the second one using also a K-means approach to simply stored some relevant points of the body shape of the subject. Moreover the classification is performed with a Convolutional-Neural-Network (CNN) in the first case and with a Fully-Connected-Neural-Network (FCNN) for the second one. Both models are capable of accurately detecting subjects and their sleeping postures. A tradeoff between the accuracy of the classification and the total amount of the data was searched. This can be usefull for some remote computation of the sleeping posture or for subject identification.**

***Index Terms*—** Principal Component Analysis, K-means, Convolutional Neural Network, Fully connected neural networks

I. Introduction

Sleeping is one of the main activity in our lives since we spend a huge part of our time on it. It's essential to physical and mental well-being and productivity. As consequence a good rest is fondamental for our daily routine. It has also been shown that sleeping and in particular the sleeping posture, affects symptoms of many diseases and plays an important role on detect them. Monitoring visually the sleeping behaviour of a person could be a very expensive operation in terms of time and human work. Moreover this needs to be performed exclusively in hospital environments where patients are required to stay overnight. A different approach can lead to solve this critical issues. The use of commercial pressure sensing mattresses can be the answer. With the tecnological improvement of these years and the diffusion of new techniques these pressure sensing arrays can now report valuable information about the subject and his position. Some technologies like cameras are very sensible to lightning changes or occlusion. On the other hand sensors work in a very simple and flexible way since they are completely independent from the environment: i.e. changing the mattress does not effects so much the sensing results. However, pressure mapping systems suffer from noisy measurements so it's important to avoid this noisy measurments to perform a correct classification. If sensors are placed on the whole area and samples are collected countinously (i.e. with a sampling rate of 1Hz) the amount of data can be very huge to transmit. This can make difficult to perform a remote monitoring over the subject and hospitalization can be preferred. The dimensionality reduction of data is then required, but this has to be done in order to maintain the correctness of predictions. Moreover a flexible model needs to be trained since it's very unlikely that someone sleeps in the exactly same way all the times. Differences can be detected between samples of the same person that stays in the same position. These are caused by change of breathing or little movements that are inevitably done during the whole sleeping period.

In this paper, we propose a bed-posture and subject classification through two different types of model according on how data were preprocessed. From the same dataset two neural networks were built for the two tasks. In other words for the dataset built with just the PCA two CNNs are used and for the dataset built with the PCA and the K-means two FCNNs are used. The goal was to obtain good performances in accuracy prediction even if the total amount of data were significantly reduced. We illustrate that both the strategies lead to acceptable results.

This report is structured as follows. In Section II we describe the state of the art, the system and data models are respectively presented in Sections III and IV. The proposed signal processing technique is detailed in Section V and its performances evaluation is carried out in Section VI. Concluding remarks are provided in Section VII.

II. Related work

At the state of art, a lot of works are done in sleep posture monitoring, because it is a very important field for clinical research, healthcare and health promotion, and also because machine learning can help posture detection and subject identification to increase the quality of this study cases.
In [1], they collected the results of a study conducted in two separate experiments using two commercial pressure mats. Moreover, they developed a deep learning algorithm for subject identification and sleeping postures identification using the pressure distributions collected. The trained models worked quiet good, but they implemented three models separately for each posture.
As shown in [2], thanks to deep learning models we are able to detect sleeping postures and subjects who are sleeping, using only the publicly available data acquired from a commercial pressure mapping system. In this paper, Davoodnia and Etemad built a deep convolutional neural networks (CNN) to classify both the subject and one of the three main sleeping positions. They considered all the pressure maps as a single frame of data, and through a loss parameter which acts as the coefficient for both classifications, they achieved the 99% of accuracy. To perform this task, they used a very large dataset which contains almost all of the samples collected by the sensors, occurring in a heavy redundant dataset, to build an enough complex model to avoid overfitting.
A very interesting approach to determine the posture of a sleeping person is shown in [3]. They made recognition based on body parts, heads, shoulders and hips, which are the maxima pressure points. This work highly underlying the main body parts of a sleeping person, probably with an additional implementation of machine learning techniques, it can achieve very good results.

III. Processing Pipeline

In this section a high level introduction of our work is presented. The main structure of the dataset and the two experiments done are briefly explained. The following picture shows the processing pipelines used:

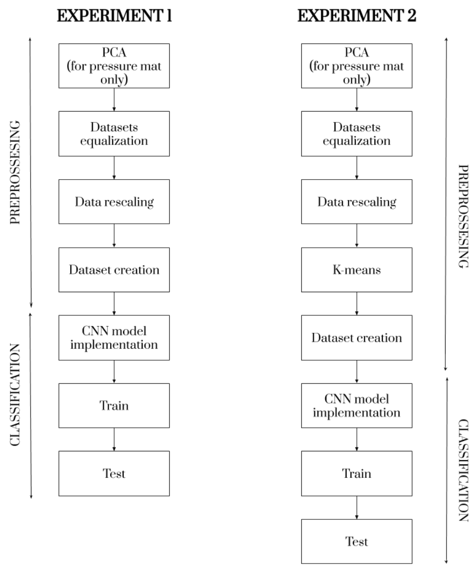

Fig. 1. Processing Pipeline for both the experiment

A. Dataset

We used pressure map data provided in a public dataset, PmatData. This dataset is composed of two different kind of collected data. The first part of data has the following characteristics:

- data is collected using Vista Medical FSA SoftFlex 2048;
- the data is collected for pressure mat;
- size of pressure mat is 64*32. This is the raw data collected reporting numbers in range of [0-1000] for each sensor;
- sampling rate is 1Hz;
- each file includes the data frames around 2-mins (around 120 frames);
- the number of subjects is 13;
- for each subject there are 17 different postures.

Some examples of the first type of pressure maps are illustrated in Figure 2.

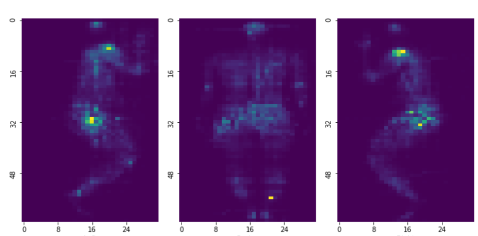

Fig. 2. First type of data.

The second part of data has the following characteristics:

- data is collected using Vista Medical BodiTrak BT3510;
- the data is collected for sponge and air mattresses separately;
- size of pressure mat is 64*27. This is the raw data collected reporting numbers in range of [0-500] for each sensor;
- sampling rate is 1Hz;
- each file contains the average of 1 frames;
- the number of subjects is 8;
- for each subject there are 29 different postures.

Some examples of the second type of pressure maps are illustrated in Figure 3.

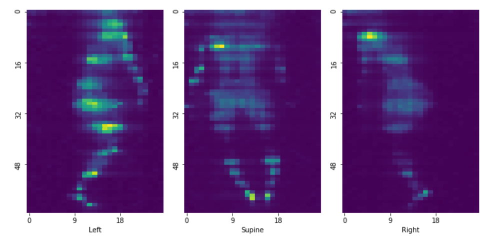

Fig. 3. Second type of data (Air matress).

The age of participants was $19-34$ years, with a height and weight of $170-186$ cm and $63-100$ Kg respectively for both the parts. In this project we combined the two parts of the data, after preprocess them, to get just one.

B. Experiment 1

The first solution proposed uses the Principal-Component-Analysis (PCA) combined with a Convolutional-Neural-Networks (CNN). According to the fact that a person could sleep in the same position for a certain time it's very inefficient to evaluate all the samples collected to perform the classification. The main idea was to try to summarise all the samples of the subject that stays in the same position in order to be able to transmit as less data as possible. So if we are able to understand when a subject is changing his/her sleeping posture then we can trainsmit just the projection of the whole samples into a new base that captures the main structure of these samples. Also the CNNs for both the classification tasks are trained with the main projections of the subject position and not all of them.

C. Experiment 2

The second solution proposed uses the Principal-Component-Analysis (PCA), like the first one, followed by a K-means approach. All of this is combined with a Fully-Connected-Neural-Network (FCNN). Also here the main idea was to try to reduce the amount of data to transmit to perform the classification. Avoiding to pass all the pressure map the K-means strategy allows us to store just the centroid positions of the subjects shape and the mean of their intensity values rispectively. As in Experiment 1 (*subsection B*) before doing that we performed a PCA. Also the FCNNs for both the classification tasks are trained with the K-means result of the main projections of the subject position.

IV. Signal and Features

As briefly described in *Section III*, *subsection A*, the PmatData dataset is composed of two different parts. Considering that we decided to use both of them and that we needed the same signals structure preprocessing was needed to perform the classification. For the pressure mat signals more samples were collected and after removing the first three, which were very noisy, a Principal-Component-Analysis (PCA) is performed. After that, according to the fact that there were some differences on both the size and intensity values of the pressure maps (64*32 / 64*27 and 0-1000 / 0-500 respectively for the first part and the second part of the PmatData dataset) a padding and a change of scale are done. Additionally, for the second Experiment, also the K-means over all the pressure maps is performed.

A. Principal-Component-Analysis

Principal Component Analysis (PCA) is a very reliable dimensionality reduction technique. It's used to extract relevant information in a big confusing dataset. The goal of PCA is to find a new basis to re-express a dataset in order to reveal interesting structure. The central idea of PCA is to find out redundancy between variables and remove it as much as possibile.
The main steps of PCA are the following:

1) Compute the data mean vector from the data input matrix X;
2) Subtract off mean vector from the dataset;
3) Calculate the sample covariance matrix Cx;
4) Calculate eigenvectors of matrix Cx in order to obtain the matrix P (new base);
5) Apply the change of base $PX = Y$.

Y will be the tranformed data matrix, if diagonal, is uncurrelated and so we won't have any redundacy.
After applying the PCA we want to reduce the dimension of the original dataset to $k < m$ where m is the dimension of the input matrix X and k the new dimension of the tranformed matrix Y. To understand how good is the approximation the *distortion measure J* is used. This can be described as:

$$J = \sum_{i=k+1}^{m} \lambda_{\mathrm{i}} \tag{1}$$

which correspond to the sum of the discarded eigenvalues, that needs to be minimized. In this project, instead of computing the *distortion measure* something similar is used which is the ratio between the eigenvalues that are kept and the whole sum of the ones of the original matrix X:

$$J = \frac{\sum_{i=0}^{k} \lambda_{\mathrm{i}}}{\sum_{i=0}^{m} \lambda_{\mathrm{i}}} \tag{2}$$

which gives us the "percent" of information lost.
In this work the PCA is used to reduce the number of samples collected from the same subject that stays in the same position. This is useful to get the position in which the subject can be in a more clear and general way. For all the positions just the 10 principal projection are kept to build the dataset.
In Figure 4 an example of the PCA result can be seen:

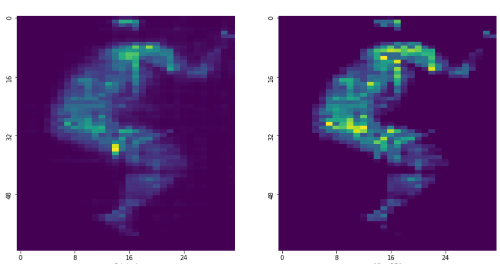

Fig. 4. Example of a PCA result

B. K-means

K-means is a clustering procedure that want to partition the points that belong to a multidimensional space into K disjoint clusters using as criterior the Euclidean space. In other words if the clusters are found then the sum of the squares of the distance of each data points to its closest centroid (center of the cluster) is a minimum.
The main steps of K-means are the following:

1) Choose some initial point for the centroids, these can be choosen at random or manually fixed;
2) Assign each data point to the closest centroid;
3) Compute the baricenters of each cluster and move the centroid in its;
4) Reiterate step 2) and 3) until a maximum number of iteration is reached or there are not significative changes in the centroids position.

In this work, the K-means procedure is used dividing the pressure map in three different sections and keeping for each point its position and its intensity value. The sections correspond to the head, the body and the leg of the subjects. For each section 2 clusters are searched in order to divide the part of the body from the matress. Moreover a mean of each cluster belonging to the body shape is computed. In this way instead of storing all the values of the pressure maps only the three centroids and thier relative intensity values are stored. This procedure was selected to reduce the amount of data significantly and check if in this way we can have a good accuracy in the classification.
In Figure 5 it is possible to see an exemple of the result of the K-means: first the centroids are shown and then the clusters bolonging to each of them.

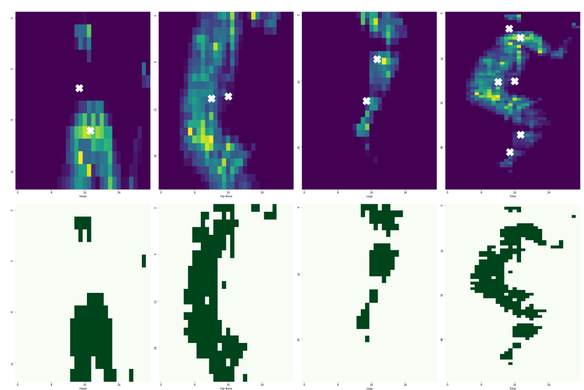

Fig. 5. Example of a K-means result. On the top there are the pressure map and the detected centroids. On the bottom there are the clusters: the green color corresponds to the subject shape and the white color to the matress.

C. Change of scale and padding

The two parts of the original dataset have as main differences the intensity range of sensors and the size of the pressure maps. To build a new unique dataset we need to fit together these two parts. First of all a change of scale is applied to make the range of the intensity values between 0 and 500. In order to preserve the ratio between the values of the same samples, a transformation is applied to each of them. This is the following:

$$NewValue = \frac{(OldValue - OldMin) * (NewMax - NewMin)}{(OldMax - OldMin)} + NewMin$$

After that a padding is applied to the pressure maps of the second part of the dataset to make them of size equal to 64*32. This padding procedure essentially adds some 0 values at the beginning and at the end of the pressure map to make it reach the desired size. Thanks to these two simply steps we are able to make a uniform dataset that cointains both the two different parts of the original one.

V. Learning Framework

For both experiments we used a machine learning approach for subject and posture identification. In first the experiment the classification is carried out with a Convolutional-Neural-Network (CNN), while in the second one we used a Fully-Connected-Neural-Network (FCNN).

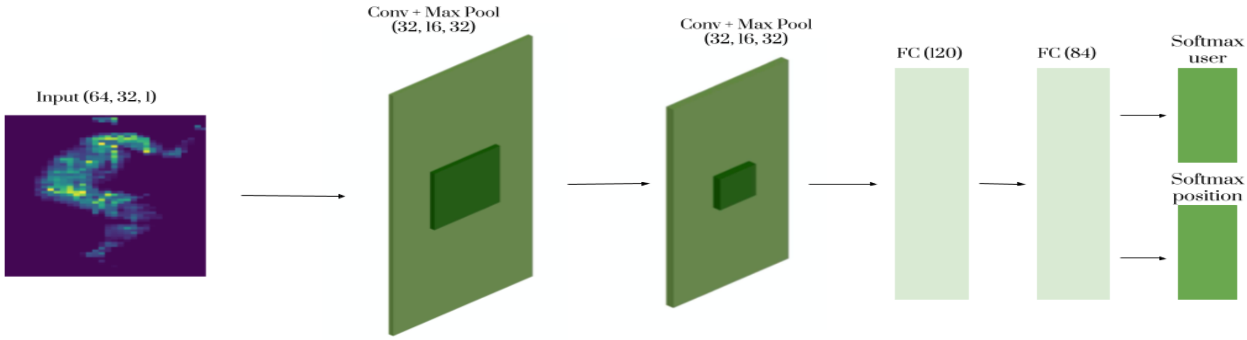

Fig. 6. Architecture of the CNN for posture classification.

A. Experiment 1

As shown in Figure 6, the classification is done using two blocks composed by a convolutional layer to detect the features, a max pooling layer to reduce the dimension and a dropout layer to regularize the data during the train to reduce overfitting. Then, it is achieved by feeding the output with two dense layers followed by two SoftMax. For the subject recognition, we added another Conv-MaxPool-Drop block to optimize the accuracy. This network's structure is used in parallel to classify subjects and postures separately. After the dataset creation, we obtain 2673 samples, which are splitted, after a data shuffle, in train set, validation set and test set. The first two contains respectively 2400 and 200 samples, and they are used for the CNN training. In the train phase of our Feed-Forward-Neural-Network we used an adam optimizer to update network weights iteratively, and we set for posture recognition 10 epochs and for subjects recognition 200 epochs because the model required more time to learn. In the subject classification training, in addition to the other parameters, we setted a batch size of 8, in this way the number of samples that will be propagated through the network are more, and the model can learn better. In the last part of the training, we used a categorical cross entropy, which is a loss function that is used in multi-class classification tasks. For the subject classification we identified 17 different subject, while in the posture recognition we had 3 labels: supine,right and left. The validation set was used to held back from training our model that is used to give an estimate of model skill while tuning models hyperparameters. This is different from the test set, because with test data we give an unbiased estimate of the skill of the final tuned model when comparing or selecting between final models. This latter is composed by 73 samples. The metrics used are the accuracy and the loss.

B. Experiment 2

As illustrated in Figure 7, the classification is done using only dropout layers, flatten layers and fully connected layers, because they are able to very effectively learn non-linear combinations of input features. Unlike the first experiment, where we labeled subject and posture in paralell, here, we used two different networks. For the subjects identification we used 5 hidden FC layers, while for the posture recognition 4. The results is achieved, in both cases, by feeding the outcome by a fully connected layers and a softmax. As in experiment 1, after the dataset creation, we obtain 2673 samples. We splitted in the same way as before in train, validation and test sets. Also in train phase of our fully connected neural network we used an adam optimizer, but we set for posture recognition 250 epochs and for subjects recognition 1000 epochs. In this alternative experiment, more epochs are needed because we iterate again for the gradient descent to converge better. As before we set a batch size of 8 and we used a categorical cross entropy. The output labels for the classification is for the posture recognition supine, right and left, while in the subject model the labels are composed by 17 subject. The metrics used are also in this case accuracy and the loss.

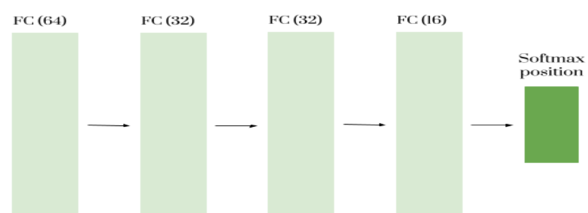

Fig. 7. Architecture of the FCNN for posture classification.

VI. Results

In this section the results are presented for both the experiments that were conducted: starting from the preprocessing result, to build a more compact and uniform dataset, ending with the accuracy metric of the tested Neural-Networks. These results are treated separately with respect to the experiments because different approaches and different models were used.

A. Experiment 1

According to the result of Principal-Component-Analysis (PCA) (Section IV, Subsection A) the number of eigenvectors that can be kept without loosing information is just one. This approach is applied just to the first part of data of the original dataset, in which there are more samples per positions. So as a consequence, if we are able to understand when a person is changing his/her position, then we can transmit just this single projection instead of all the measurments taken to perform the classification.

The dataset was built using the top ten principal components obtained through the PCA for each posture. Moreover, the change of scale and a padding, where required, are applied to all the samples. The total size of the data is reduce from 106MB (size of the original one) to 87MB for train, evaluate and test the models. The two Convolutional-Neural-Networks used for subject and posture classification lead to an average accuracy of 93% and 97% respectively.

In Figure 8 the accuracy increasement over the epochs of both the models is shown:

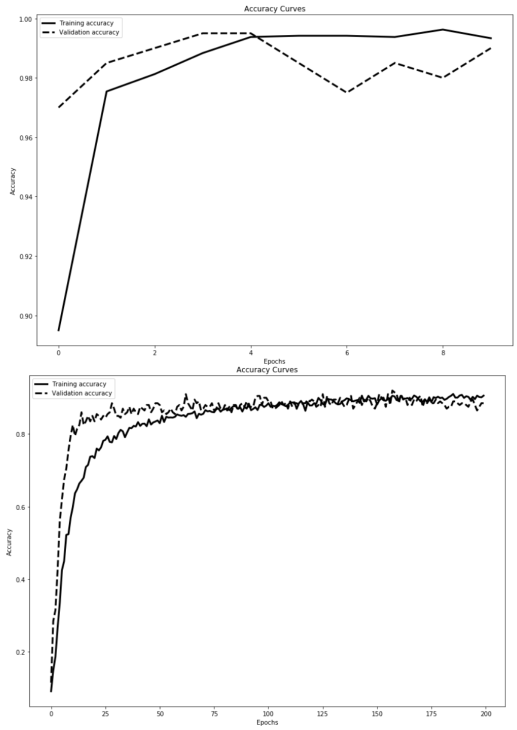

Fig. 8. Results of the accuracy over the epochs. On the top the accuracy result of the CNN for position classification. On the bottom the accuracy result of the CNN for subject classification.

As reported above the accuracy of the two models reaches good values and the PCA results reduce the amount of data needed to be transimitted. Once again, if a patient has been monitoring during his/her sleeping period, just the PCA result needs to be passed to perform a correct classification.
In Figure 9 the confusion matrix of position and subject classification are shown:

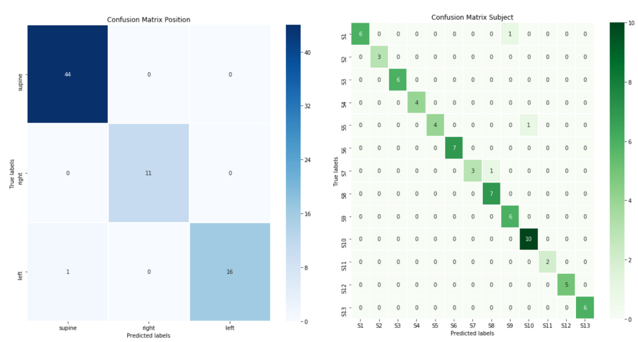

Fig. 9. Confusion matrices. On the left the one for position calssification and on the right the one of subject classification.

B. Experiment 2

As for Experiment 1, the Principal-Component-Analysis (PCA) is used over the first part of the original dataset. In this case the top ten principal components for each position are kept. Moreover the change of scale and a padding, where required, are applied to all the samples. After that the K-means approach (Section IV, Subsection B) is applied on all the pressure maps. To build the new dataset just the three centroids positions and the mean value of their clusters are stored. The total size of the data is reduced from 106MB (size of the original one) to 399,9kB for train, evaluate and test the models. The two Fully-Connected-Neural-Networks used for subject and posture classification lead to an average accuracy of 70% and 95% rispectively.
In Figure 10 the accuracy increasement over the epochs of both the models is shown:

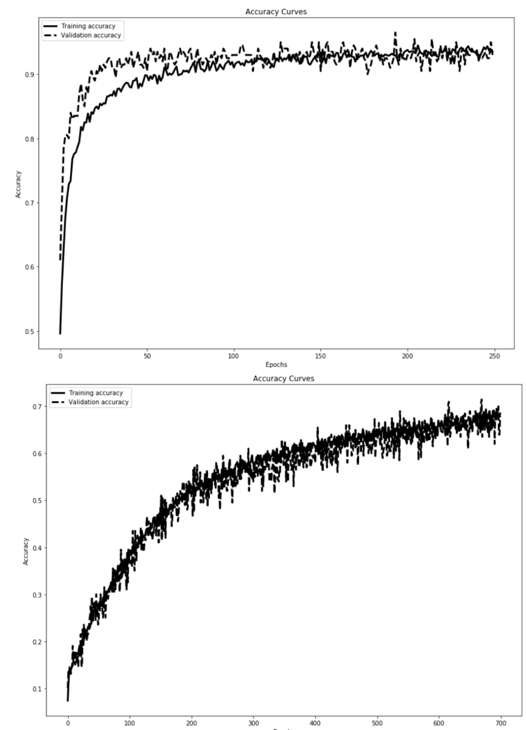

Fig. 10. Results of the accuracy over the epochs. On the top the accuracy result of the FCNN for position classification. On the bottom the accuracy result of the FCNN for subject classification.

As reported above the accuracy of the two models reach good values too, with respect of the use of the K-means approach. The total amount of data is drastically reduced: instead of passing the whole pressure map for the classification just the three centroids position and the mean values of their relative clusters are needed to perform a correct classification.
In Figure 11 the confusion matrix of position and subject classification are shown:

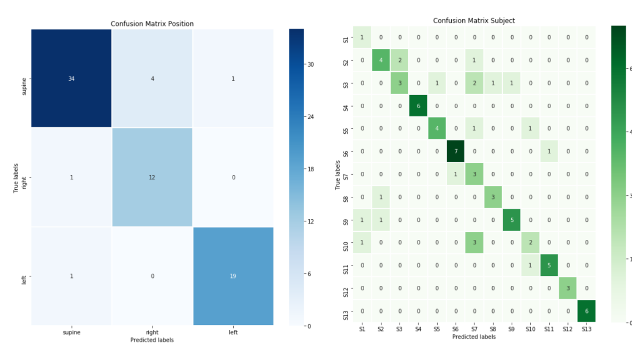

Fig. 11. Confusion matrices. On the left the one for position calssification and on the right the one of subject classification.

VII. Conclusion

This work wants to solve a classification problem about sleeping postures. The models implemented can detect the subject and his/her posture. The postures are divided in supine, right and left. A different neural network is built to perform subject and posture classification respectively. Moreover, different kinds of networks are implemented for the two different experiments that were conducted, as explained in Section V. According to the results presented in Section VI some considerations are needed. The classification task is reached for both the experimenst but with some differences. The experiment 1 leads to better results with respect to the models accuracy since it reached over the 90% of accuracy for both the subject and posture classification. On the other hand the dataset used to train and test the networks has a lower size in respect to the original one but not so significantly. In a real monitoring situation, asssuming that we can perform a PCA over all the samples of the same position and just one pressure map can be passed, we have in any case to pass 2048 intensity values. The experiment 2 leads to a high accuracy in postures clasification but a lower one to the subject classification. On the other hand the amount of data is drastically reduced thanks to the K-means approach. So in a real monitoring situation, asssuming that we can perform a PCA over all the samples of the same position and just one pressure map can be passed, we have to transmit just the centroids postions and the mean of the intensity values of their relative clusters. The lower accuracy in the subject identification can be explained thinking about the way in which the information about intensity values is obtained. With the mean of the clusters intensity values we are approximating the distribution of these. But if the clusters follow a certain distribution we would be able to make easier to our network to differentiate the subjects. Moreover passing just the intensity values to the Fully-Connected-Neural-Network instead of passing also the centroids positions, bring to the loss of the possible information of the subjects height that can be discriminated. This was avoided by the fact that the centroids positions were not so precise to bring this kind of information correctly to the FCNN.

A. Personal difficulties

In this project the main difficulty that we have encountered was related to the original dataset. There were differences in what was reported in the text file containing the structure and organization of the data and how effectively this was. Assuming that the structure and the organization reported were reliable we encountered some errors during the programming phase. To solve these issues we also needed to change name to some files and in another case we needed to remove some of these.
The main differences that we encountered were:

- it was written that for the second part of the dataset "each file contains the average of around 20 frames" instead just one single frame for each position was present;
- it was written that the subjects selected for the second part of the dataset were from 1 to 8 and 12, instead the last one is missing and the real subjects were 1,2,3,5,7,8,9,10;
- the positions of the second part of the dataset were identified with letter from B to F but for just one subject there were also some G file which were not reported;
- the position were mirrored, so for the first part of the dataset a right posture looks like a left one and viceversa;
-

References

[1] M. Baran Pouyan, J. Birjandtalab, M. Heydarzadeh, M. Nourani, *A Pressure Map Dataset for Posture and Subject Analytics*. Quality of Life Technology Laboratory The University of Texas at Dallas, Richardson, TX, 75080 USA.

[2] Vandad Davoodnia and Ali Etemad, *Identity and Posture Recognition in Smart Beds with Deep Multitask Learning*. 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) Bari, Italy. October 6-9, 2019.

[3] Tae-Hwan Kim, Soon-Ju Kwon, Hyun-Min Choi and Youn-Sik Hong, *Determination of Lying Posture through Recognition of Multitier Body Parts*. 2019 Hindawi Wireless Communications and Mobile Computing.