

Subject and Posture Classification with CNN

Gaurav Anand

gaurav.anand@studenti.unipd.it

Abstract—Sleeping is one of the most important activities in our daily life that affects our health. However, few people really know their sleep habits, which is important to avoid possible sleep-related diseases. One benefit of sleep monitoring is that it can lead to positive changes. People are more likely to change their habits if they track them. Pressure sensor mats consist of grid-like and flexible force sensors that are now commercially available to continuously measure the pressure distribution under body parts in different sleeping positions. In this study, we propose a convolutional neural network (CNN) for three different classification tasks. The first task is capable of accurately detecting subjects, the second task is capable of detecting three standard sleeping postures (supine, right and left) and the last one is a multitask classification of subject identification and posture recognition at the same time. We evaluate the performance of our models applying it on two different data-sets. Our model showed really promising results for both experiments used in each classification tasks.

I. INTRODUCTION

Sleep plays a crucial role in health and well-being in life. Getting enough quality sleep at the right time can help protect your mental and physical health, quality of life, and safety. The pattern of posture in bed during sleep is one of the most important indicators of sleep quality and can be useful for various medical diagnoses.

Recently, sleep posture has been shown to be important in monitoring health conditions such as congestive heart failure (CHF), pressure ulcers, and even blood pressure abnormalities. Recent studies have shown that posture in bed has a major impact on the occurrence of sleep disorders such as apnea. The conventional system for studying sleep using neurophysiological signals, polysomnography (PSG), is extremely complex and expensive. In addition, many PSGs are performed exclusively in hospitals, requiring patients to stay in the hospital overnight.

In this paper, we present three different classification tasks using a convolutional neural network to easily recognise subjects and sleep postures from a dataset. Considering two different datasets of pressure sensor containing different subjects and different sleeping postures, the following classification tasks were implemented: Subject Identification, Posture Detection, and Joint Subject Identification and Posture Detection, where for each of these tasks we compare the results of the two presented datasets. Researchers have also found that sleep quality is related to sleep position and that frequent postural changes during sleep, such as snoring or extensive body movements, can lead to shorter sleep duration. Recently, pressure measurement mats have become commercially available that continuously measure the distribution of pressure under the body while lying in bed. These pressure mats provide very

valuable data that can be used directly by machine learning algorithms to track posture and assess the stress on individual body parts. Our method can provide valuable information in a faster way that can be used in smart home and clinical settings, where the information obtained can be relevant to the association with certain diseases.

This report is organised as follows. In Section II we describe the state of the art. Whereas regarding dataset, pre-processing techniques, learning framework and results are presented in Sections III, IV, V and VI respectively. The proposed signal processing technique is described in Section IV and its performance is evaluated in Section VI. Concluding remarks are given in Section VII.

II. RELATED WORK

Over the years, many different methods of assessing sleep have emerged, especially in recent years with new methods and models. Therefore, monitoring and relieving the pressure generated by various postures is a critical step in reducing the incidence of pressure ulcers in hospitals and nursing homes.

Interest in sleep quality assessment has increased, not only in clinical research and treatment of sleep disorders, but also in health care and health promotion. However, unlike the data-set used in our study, the number of postural classes in previous studies has been limited. Studies on the effects of sleep postures on sleep apnea and ulcer prevention using advanced monitoring technologies for smart bed applications have also gained attention. This work has greatly helped us understand how our study can be applied in the real world, but the methods suffer from occlusions, brightening conditions, and viewing angle and calibration effects.

Another study showed that assessment of posture in bed is important in evaluating sleep quality and health status of bed-ridden patients. They used one of the datasets we considered in our work, but instead of working with raw data as we did, they extracted 18 statistical features from the print data.

Another interesting work is the use of a deep end-to-end CNN framework proposed for joint classification of subject identification and body posture detection with pressure sensors in bed. Their work is an efficient and highly accurate model that performs both tasks simultaneously, namely subject and posture recognition, where conventional classification algorithms do not achieve good accuracy on unseen data. They have shown that their method significantly outperforms other work on this study.

Their proposed algorithm can eventually be used in clinical and smart home environments as a complementary tool to other available automated patient monitoring systems.

III. DATASET

In this study, we used a public dataset, PmatData, which contains in-bed postural pressure data from multiple adult participants who used two different types of pressure sensing mats. All 13 participants were healthy individuals with no sleep disorders or history of pressure ulcers. Individuals 1-8 participated in both experiments, while individuals 9-13 participated only in Experiment-I. PmatData contains pressure data from two separate experiments. These two experiments are:

- Experiment-I: The first dataset contains pressure data from 13 participants in 8 standard postures and additional conditions. The size of the pressure mat is 32*64. This is the raw data collected, reporting numbers in the range of [0-10,000] for each sensor. The sampling rate is 1 Hz. Each file contains data frames of about 2 minutes (about 120 frames). For this data set, we have an input field of the form. (20024, 2048) containing all the pressure mats, and two labeled arrays of shape (20024,), one containing the subject labels and the other the posture labels for each pressure mat. The labels for subject recognition in this dataset are all subjects from participant 1 to 13 and for posture recognition we have three standard classes: Supine -Left -Right.
- Experiment- II: The second dataset contains pressure data from 8 participants in 29 different conditions of 3 standard postures. In addition, the experiment II was collected separately for normal and air pressure mattresses. The size of the pressure mat is 27*64. This is the raw data collected reporting numbers in the range of [0-500] for each sensor. The sampling rate is 1 Hz. Each file contains the average of about 20 images. The data is collected separately for both sponge and air mattresses. For this data set, we create an input array of the form (462, 1728) and two labeled arrays of shape (462,), one containing the subject labels and the other the posture labels for each pressure mat. Here, the subject labels are only for the first 8 participants, while the posture labels are the same three standard classes: Supine- Left-Right.

For Experiment-i, each file contains multiple frames of recordings, where the 2048 columns represent the 32x64 pressure sensor measurements and each row contains a new set of samples. For experiment-ii, each file contains an average of multiple frames of recordings. Unlike experiment-i text files, the 27 columns and 64 rows represent the 27x64 sensor matrix.

IV. PROCESSING PIPELINE

In our project, we implemented a similar convolutional neural network for each of the classification tasks. The architecture of the models differs only in the output layer.

After trying different implementations for the architecture of our models, we take inspiration from the paper. Each classification is performed by a feed-forward pass, which consists of two blocks composed of the following layers: Convolution layer, BatchNormalization layer, MaxPool layer, and

LeakyRelu layer. This is followed by two more blocks with the structure: convolution layer, BatchNormalization layer and LeakyRelu layer. These four blocks capture the important properties of each frame.

Classification is then performed by feeding the results into a dense layer. For subject identification and posture recognition, the dense layer is followed by a SoftMax layer. In contrast, for joint identification of subject and posture recognition, the dense layer is followed by two SoftMax layers, which are used in parallel to classify subject and posture separately.

We used the Dropout technique to prevent overfitting. Below in Figure 1 is the architecture of our model.

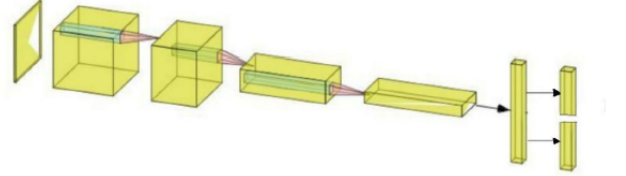


Fig. 1: Our proposed framework is presented

In the pre-processing step we attempted to reduce the noise caused by occasional malfunctioning pressure sensors. Such artifacts often take place when individual sensors are subjected to large pressure values outside of the allowed voltage range.

- To cope with this problem, we used for both data-sets the Median Filter of size 3x3x3 pixels, we tried first a filter size of 1x3x1 but the image was not as clear as the 3x3x3 size. The main idea of the median filter is to run through the signal entry by entry, sort them and replacing each entry with the median of neighboring entries.
- Second thing we did, only for the first experiment, was eliminating all the images that were not clear enough (most of the values of raw data were 0).
- And to conclude with the pre-processing we normalized the input data by dividing each value with the maximum range that a sensor could get, respectively for each data-set, we divided by 10,000 (range for each sensor was from 0- 10,000) and by 500 since the second data-set had its range for each sensor varying from 0-500.
- Before implementing the model, we divided both data-sets into three parts: train, test and validation sets; with a test size of 0,25. Also, for experiment-ii, since we have less data, to prevent overfitting we increased the number of the dataset.

V. LEARNING FRAMEWORK

The models proposed were implemented using the Tensor-Flow platform. After applying different models, we selected the model that performed the best and provided the best results. As mentioned in the previous section, we used a convolutional neural network as the final model for three of the classification tasks presented. In the following points, we describe the architecture of our model in more detail.

- First Block: Conv2D layer of filter size 32, kernel size (3,3) with a stride of 1, a Batch-Normalization layer, MaxPooling layer of kernel size (3,3), and a LeakyReLU layer with an activation coefficient of 0.2. After the LeakyReLU layer, we added a dropout layer with a rate of 10% to prevent overfitting.
- Second Block: It again has exactly the same levels as the first block, but here we change the filter size of Conv2D from 32 to 64 and the dropout rate to 20
- Third Block: It consists of a Conv2D layer, a Batch-Normalization layer, and a LeakyRelu layer. Starting from this block, we removed max-pooling and also changed the filter size of Conv2D from 64 to 128 and the dropout rate to 30%.
- Fourth Block: This block contains the same levels as the third block, but we have increased the filter size from 128 to 256 from Conv2D and also increased the dropout rate to 40%.
- Fifth Block: In the last block we have a total of different layers. We used Flatten layer that receives the output of the convolutional layers to create a single long feature vector, a dropout of 50% and Dense layers with a softmax activation and a size of 256 and 14 or 4 depending on the classification task: subject or posture, or for the multitasking problem where we use two parallel softmax Dense.

We trained our convolutional neural network using the Adam optimizer with a fixed learning rate of 0.00002. Adam is an adaptive learning rate optimization algorithm specifically designed for training deep neural networks. The loss function used in the model is the categorical cross entropy, which is used for multi-class classifications. This loss function is also called Softmax Loss and is composed of a Softmax activation plus a Cross-Entropy loss. Using this loss function, we train a CNN to output a probability over the C classes for each image. In the special (and common) case of Multi-Class classification, the labels are one-hot, so only the positive class C_p retains its term in the loss function. Figure 2 shows a diagram of the loss function used.

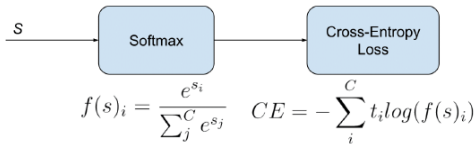


Fig. 2: Categorical Cross-Entropy loss

There is only one element of the target vector t that is not zero $t_i = t_p$. Thus, if we discard the elements of the summation that are zero because of the target labels, we can write:

$$CE \equiv - \log \left(\frac{e^{s_p}}{\sum_j^C e^{s_j}} \right)$$

where s_p is the CNN score for the positive class.

For the first two classification tasks, subject identification and posture recognition, only a categorical cross-entropy loss function is generally used. For the multitask classification of joint subject identification and posture recognition, we use two categorical cross-entropy loss functions (as we implement the model for classifying at the same time subject and posture recognition).

VI. RESULTS

As explained in the previous sections, we used a Deep Learning approach, i.e., a convolutional neural network model, for our three classification tasks: Subject Identification, Posture Recognition, and Joint Subject and Posture Recognition. We used three different metrics for the evaluation:

- Accuracy: it is the fraction of predictions our model got right
- F1 score: it is the harmonic mean of the precision and recall
- Confusion Matrix: is a table often used to describe the performance of a classification model (or "classifier") on a set of test data for which the true values are known.

Let us first consider the subject identification task, where we need to identify each participant for both data frames. The Convolutional Neural Network model achieved high accuracy (99.4%) in the first experiment, although it is a bit overfitted, as we can see in Figure 3.

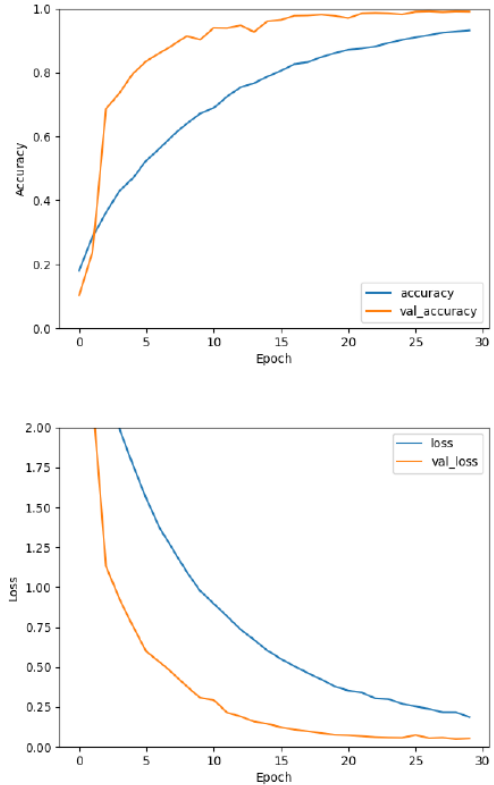


Fig. 3: Accuracy and loss for subject identification

After this analysis, we checked the performance of the model against the dataset of Experiment-ii and we obtained more or less the same behaviour for the overfitting, instead the values of the evaluation metrics are quite different, as shown in Table 1.

Metric	Experiment i	Experiment ii
Accuracy	99.4%	83.9%
F1 score	99.3%	82.5%

TABLE 1: Subject identification Results

Based on these results, we can conclude that the model performs better in the first experiment, which we believe is due to the fact that the second experiment has fewer participants and fewer frames are recorded per participant. The performance of our model can be also observed in the following Confusion Matrix showed in Figure 4.

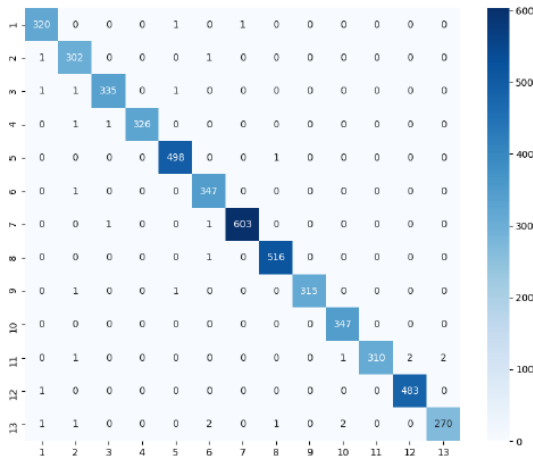


Fig. 4: Confusion Matrix subject identification

Let us now focus on the classification task of posture recognition. The performance obtained with the Experiment-i dataset in posture recognition is better than in subject identification. In Figure 5 below, we can see that there is no overfitting effect.

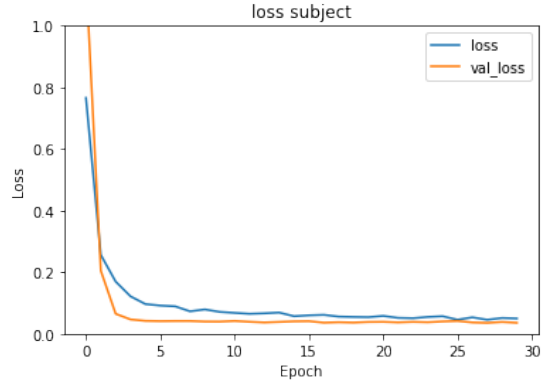
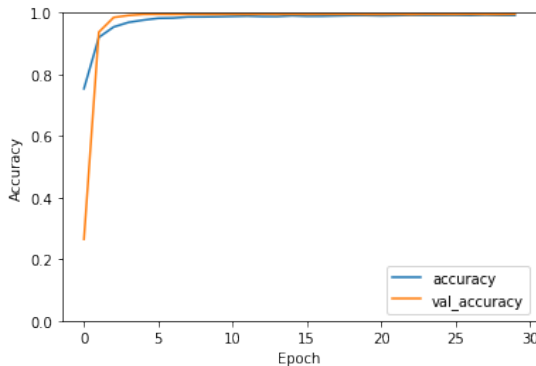


Fig. 5: Accuracy and loss posture recognition

We achieved an accuracy of 99.3% and an f1 score of 99.3%. Figure 6 shows the confusion Matrix performance. When we compare the results of the first and the second experiment for this classification task, we notice that the performance of the model decreases in the second experiment, as we can see in Table 2. Thus, the model behaves in the same way as it does for subject identification.

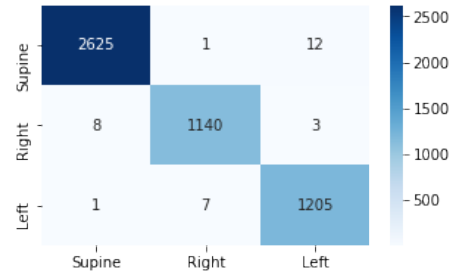


Fig. 6: Confusion Matrix posture recognition

Metric	Experiment i	Experiment ii
Accuracy	99.3%	74.9%
F1 score	99.3%	75.4%

TABLE 2: Posture recognition results

As the final step the focus was on the analysis of the last classification task i.e. joint subject identification and posture recognition. Since we implemented a multi-tasking learning for this classification, we now show in the tables (3,4) the performance for each task on each data-set.

Metric	Subject Id.	Posture Rec.
Accuracy	99.36%	99.32%
F1 score	99.36%	99.32%

TABLE 3: Experiment-i: Multitasking learning results

For experiment-i, if we compare the results of the multi-tasking learning with results obtained in the first two classification task we observed similar performance. Instead, for the

Metric	Subject Id.	Posture Rec.
Accuracy	80.27%	79.06%
F1 score	79.78%	77.62%

TABLE 4: Experiment-ii: Multitasking learning results

experiment-ii the results in the multitasking learning decrease. In our opinion this different behaviour between the two experiments is due to the fact that the two data-sets are very different.

VII. CONCLUSIONS

To conclude, the convolutional neural network model was implemented for the three different classification tasks mentioned in the previous section. Our work is efficient and very accurate. It can be of great use in clinical and smart home environments as a complementary tool. Future work can focus on pressure datasets with higher temporal resolution to measure physiological signals such as heart rate.

In this project, it was a bit of struggled to find the best model because before deciding on the final model, several models were implemented and verified that they worked well with the datasets from both experiments. To summarize what we learned from our work in the following points.

- Importance of sleeping posture;
- Working with sensors data;
- Pre-processing techniques;
- Working with two different data-sets at the same time;
- Multitasking learning;
- Implementing robust CNN models;
- How to handle a classification problem in real life.

REFERENCES

- [1] M. B. Pouyan, J. Birjandtalab, M. Nourani, and M. M. Pompeo, "Automatic limb identification and sleeping parameters assessment for pressure ulcer prevention," *Computers in biology and medicine*, vol. 75, pp. 98-108, 2016.
- [2] S. Ostadabbas, R. Yousefi, M. Nourani, M. Faezipour, L. Tamil, and M. Q. Pompeo, "A resource-efficient planning for pressure ulcer prevention," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 16, no. 6, pp. 1265-1273, 2012.
- [3] M. L. Blood, R. L. Sack, D. C. Percy, and J. C. Pen, "A comparison of sleep detection by wrist actigraphy, behavioral response, and polysomnography," in *Sleep*, vol. 20, no. 6, pp. 388-395, 1997.
- [4] Moein Enayati, Marjorie Skubic, James M Keller and Mihail Popescu, "Sleep Posture Classification Using Bed Sensor Data and Neural Networks", in 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). July 2018.
- [5] C.H.Lee, D.K.Kim, S.Y.Kim, C.S.Rhee, and T.B.Won, "Changes in site of obstruction in obstructive sleep apnea patients according to sleep position" in *The Laryngoscope*, vol. 125, no. 1, pp. 248-254, 2015.

[6] Hao T, Xing G, Zhou G, "Unobtrusive Sleep Quality Monitoring using Smartphones" in: *Proceedings of the 8th ACM conference on (SenSys)*, 2013.

[7] R. D. Cartwright, "Effect of sleep position on sleep apnea severity," *Sleep*, vol. 7, no. 2, pp. 110-114, 1984.

[8] Vandad Davoodnia and Ali Etemad, "Identity and Posture Recognition in Smart Beds with Deep Multitask Learning", in: *IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019.