Eye-Tracker: ML Approach for SNA Activation Monitoring

Andrea Benedetti, Silvia Leccabue
University of Illinois at Chicago,
{abened6, slecca2}@uic.edu

***Abstract*— This work is inserted in the scope of the technology for ocular tracking, a tech that is used to record the dimension and the movement of the pupils.**
The principal aims are basically two: the comparative analysis of the signal recorded by two different devices, the fixed and the mobile one, and the correlation between the autonomic nervous system (SNA) activity and the pupil dilatation signal due to the stimulus induced by the IAPS images.
Pupil dilatation is known to be directly correlated to the level of light exposition of the eye and to the activity of the SNA, particularly it is controlled by the action of parasympathetic and sympathetic systems
Thanks to the precision of the newest devices for oculometry, and with the use of proper software for data manipulation, it is very interesting to find out if the pupil signal could be interpreted as a driver for the SNA.
Inspecting the correlation between pupil dilatation and SNA activity the aim was to corroborate the hypothesis shared in literature.
The final goal for this approach would be a ML algorithm able to precept in real time the SNA behaviour in the detected signal of the subject, providing a metrics for the Arousal explicated by the subject itself according to different stimuli.
In our project, we implemented an offline analysis pipeline to inspect the significance of the data and a preliminary training of the model created on inter-subject recordings.

I. Introduction

A. SNA and Pupil Diameter

The proper use of eye-tracking technologies offers the possibility to implement brain-to-computer interfaces, making the research of an application or a content more intuitive.
The basic idea under this approach is the analysis of the frequency content provided by the pupillometric data (top graph in the Fig.1), which have shown high correlation values with the HRV (Heart rate variability) signals (bottom graph in the Fig.1), the state of the art for monitoring SNA (autonomic nervous system) activity through its frequency content. [1]
The main features of interest for a discriminative approach between different emotional status linked to the arousal of the subject are circumscribed into two bands of interest, 0.03-0.2 Hz that monitors sympathetic activity and 0.2-0.4 Hz, which related to sympathetic activity.
The top graph of the Fig.1 is related to the signal acquired through Eye tracker device. The higher is the arousal and the higher is the the reached frequency, taking into consideration the LF (Low Frequency) band. On the other hand, in the HF (High Frequency) band, the frequency peak decreases as increases the arousal. [2]
Same behaviour can be observed in the bottom graph that describes the HRV signal, currently the gold standard in this specific application.

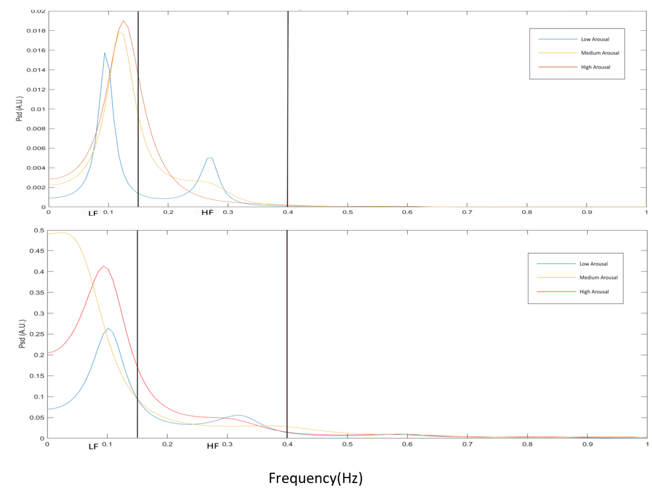


Fig. 1. The top and the bottom graphs show pupillometric signal and HRV signal respectively. Their relative trends in both the LF band and the HF band considering Low Arousal, Medium Arousal and High Arousal can be observed in both graphs.

B. Experimental Protocol

In order to create difference arousal response in the subject, we implemented an experimental pipeline where we asked to each subject to sit in front of a screen integrated with an eye-tracker system to record the data of interest.
For our visual stimuli, we used IAPS images (International Affective Picture System), a set of labelled images used for research purpose in psychological inspections. In particular, we selected images that create difference level of arousal in subjects. [3]
An example of the arousal level on which the images are labeled is shown in Fig. 2. The protocol followed during the

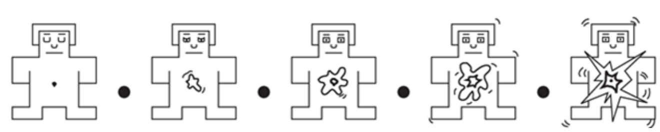


Fig. 2. Labels corresponding to a different arousal level.

data acquisition can be summarize in few main steps. Firstly, we recorded the signal from 30 subjects in which has been presented a baseline image for 90 seconds.
The first group of images undergone to each patient was the 30 images with low arousal, then 30 images with medium arousal and lastly 30 images with high arousal, as it has been summarized in Fig. 3

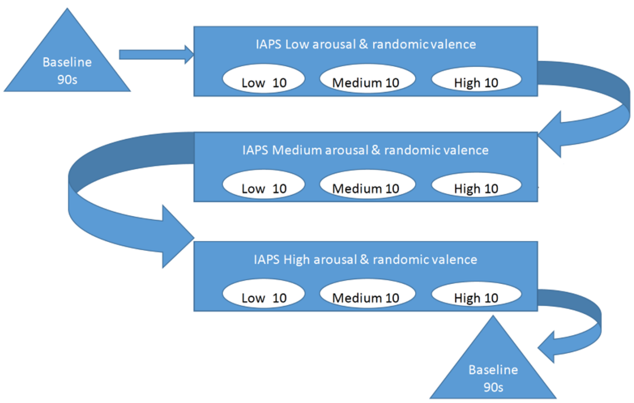


Fig. 3. Scheme of the protocol followed during the data acquisition.

II. Data Preparation

The main step proposed in this work from a data-analysis to a machine learning approach was how the data were inspected and subsequently analyzed.
From a Time-Series analysis we needed a partitioning of a Time-series per subject to matrix involving samples and features. Therefore we pass from a single vector having the three labels for each time points, to three blocks one for each label, as it has been represented in the Fig. 4.

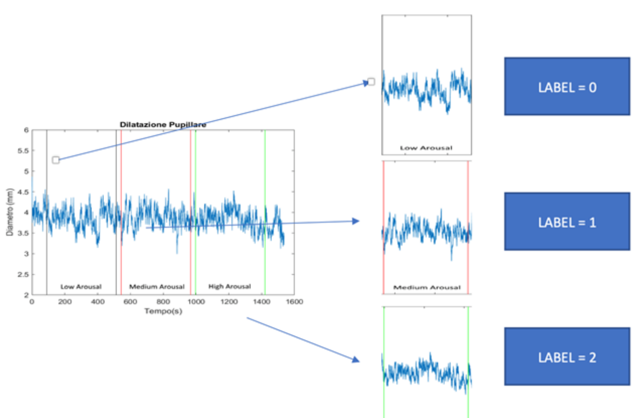


Fig. 4. Dataset preparation obtained from the portioning of the single vector in the time domain in the time domain

A. Block Processing

First of all we divided the Time-series into their respective labels according to the intensity of arousal in which they were collected; from a 29 vectors of data (one for each subject) we crop the labelled blocks into different samples creating a 84 row matrix containing 24000 columns (60 frequency sampling x 400 sec in each block). [4]

B. Feature Extraction

We performed a feature extraction based on our previous statistical analysis (described in III) and after this we decided to compute 7 features for each sample, summarized as follow:

- Sample Mean: the pupil diameter is strictly related to the Autonomic Nervous System activity and even if there is a variability between subjects, it could still be a marker for classification purpose.
- Sample Variance: the sample variance around the sample mean is often used in statistical studies as the ANOVA. [5]
- Spectral Content: we used the YULE-WALKER algorithm that computes the spectral power calculation the coefficients of the respective AR model. In particular we inspected the low frequency, high frequency, low frequency/high frequency area under the spectral curve.

III. Data Inspection-Preparation

An important consideration is that the extracted features may be measured at different scales and can be affected by the intra-patient characteristics. Therefore, we need to consider all these features could not contribute equally to the model fitting and to the model learned function and might end up creating a bias. [6]
Thus, to deal with this potential problem feature-wise, a standardization step has been performed before to model implementation and PCA analysis described in IV and V.
The main idea is to standardize each feature by subtracting its expected value μ and dividing the difference by its standard deviation σ, assuming that the dataset has a Gaussian distribution with μ = 0 and σ = 1. [7]
Moreover, the outliers analysis has been performed because the detection of anomaly is fundamental for dataset consistency and integrity. As we can observe in the following graphs in Fig.5, the population is grouped together and the outliers are plotted as single points. The technique used to identified outliers and remove them from the dataset has been the Z-score.
By definition, the Z-score is the signed number of standard deviations by which the value of a data point is above the mean value of what is being measured.
The intuition behind Z-score is to describe any data point by finding their relationship with the Standard Deviation and Mean of the group of data points. Z-score is finding the distribution of data where mean is 0 and standard deviation is

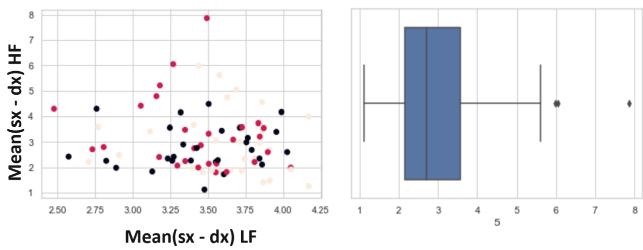


Fig. 5. On the left the scatter plot of the features Mean(sx - dx)LF and Mean(sx - dx)HF; on the right the boxplot of the outliers of the fifth feature of the dataset.

1 i.e. normal distribution. [8]
Mathematically, the formula is as follow:

$$Z = \frac{x - \mu}{\sigma} \tag{1}$$

where x is the feature value, μ is the mean and σ is the standard deviation. Instead, as regards the ML algorithm, we have used Z-score function defined in scipy library to detect the outliers before training the model.

A. Statistical Validation

Before applying a ML (Machine Learning) approach to our dataset, was performed a statistical analysis on the features of interest, in particular a Friedman test was evaluated on different metrics of the dataset.
Both in the time domain and in the frequency domain we tried to analyse the control on the pupil diameter signal of the SNA in response to different visual emotive stimuli.
In the time domain we divided the temporal signal in three band of interest (Low Arousal, Medium Arousal, High Arousal). [9]
From the Friedman tests executed we could correlate significant variations between the data we recorded, in particular we inspected the two activation bands of the SNA, one around 0 and 0.15 Hz and the (low frequencies), and the other around between 0.15 and 0.4 Hz (high Frequencies).
We obtained different results: the pupillary activity reacts to the IAPS stimuli, supporting the starting hypothesis and introducing the possibility of a direct relationship between the analysed activity and the Autonomic Nervous system.
Friedman tests reveal that the power content in the LF (sympathetic area) band grows with a significant manner from low to middle Arousal, attesting an increased activity of the autonomic nervous system. [10]
In the HF bands the opposite happens, as the Arousal grown in the subjects increase the power intensity linked to the resting state (HF, parasympathetic) shows a decrease justify the transition of the power spectra in the LF band.

IV. SVM Model Implementation

We divided the dataset into all the possible combination of labels (low vs medium, medium vs high, low vs high Arousal) to let a more intuitive interpretability of the decision boundaries during visualization. For our classification, we implemented a Gridsearch able to explore different parameters as:

- Kernel: since the lack of linearity in our data we looked for different kernels such as polynomial and rbf (radial basis function). Mathematically the polynomial kernel is expressed by the following equation:

$$k(\mathbf{x_i}, \mathbf{x_j}) = (\mathbf{x_i} \cdot \mathbf{x_j} + 1)^d \tag{2}$$

  Instead, the rbf kernel is described by (3):

$$k(\mathbf{x_i}, \mathbf{x_j}) = exp(-\gamma \parallel \mathbf{x_i} \cdot \mathbf{x_j} \parallel^2),\ for\ \gamma > 0 \tag{3}$$

- Gamma : Intuitively, the gamma parameter defines how far the influence of a single training example reaches, with low values meaning 'far' and high values meaning 'close'. The gamma parameters can be seen as the inverse of the radius of influence of samples selected by the model as support vectors.
- C: The C parameter tells the SVM optimization how much you want to avoid missclassifying each training example.

V. Results

The presented results regard the 1vs1 classification between:

- label = 0 vs label = 1
- label = 1 vs label = 2
- label = 0 vs label = 2

The best performance is obtained in the classification between low and high arousal, in accordance to the confusion matrices (Fig. 6) we can see that the accuracy level is balanced through the classes and the accuracy level of 0.88 results to be meaningful.

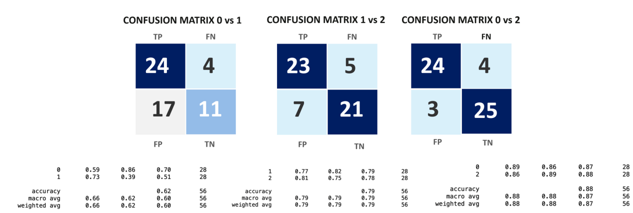
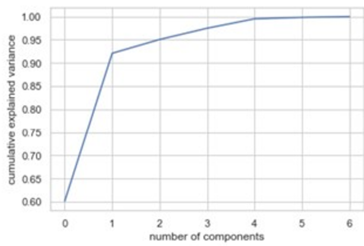


Fig. 6. On the top, the confusion matrix of each 1vs1 classification task. On the bottom the values of accuracy, macro average and weighted averaged reached by the model.

A. Interpretability

To provide interpretability of the dataset we computed a Principal components analysis in the dataset and computed a Logistic Regression on the main 2 principal components obtained as a linear combination of the starting features.
The first image in Fig. 7 shows how the heaviest components used to obtain the PC1 and PC2, which are the components that explain >90% in our dataset, are the LF/HF for the first principal component and LF and HF for the second. The values



	pc1	pc2
Mean sx	-0.012481	-0.004293
Mean dx	-0.007819	-0.014628
Var sx	0.025305	0.011705
Var dx	0.023646	-0.001368
Mean(sx - dx) LF	0.178881	0.743768
Mean(sx - dx) HF	0.152343	0.625282
LF/HF	0.971276	-0.235500

Fig. 7. On the left the explained variance ration with respect to the number of components after the PCA analysis of the data. On the right, the values of the first two principal components for each feature.

of the different values of the first two principal components are shown in Fig. 7.

Instead, the image in Fig. 8 represents the decision rule made by a SVM algorithm that is able to discriminate most of the labels associated to 0 and 2 label.

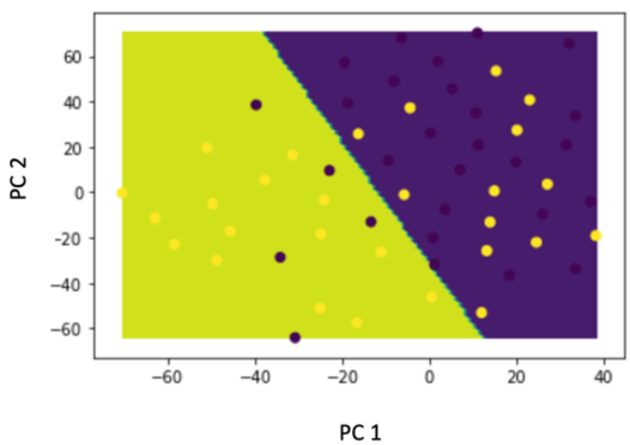


Fig. 8. Decision boundary of the labels associated to 0 and 2.

VI. Conclusions and discussions

We can conclude stating that our results open the way to innovative approach in Brain Computer Interface Design; the approach provided results in a quite good accuracy in discriminating the emotional state of the subject based only on its pupillometric data, which can bee thought as one of the less invasive way to extract information about the Autonomic Nervous System, even from remote.
Advancements in computer vision will surely provide even more promising results in BCI systems based on the analysis of the eye of a subject. This type of approaches can be used in several application, from rehabilitation to the marketing, passing through innovative way to control our electronical devices.
For future developments much still need to be experimented, first of all the experimental set-up should be designed in a way that is able to arouse an emotional response in a more immersive way than just presenting some passive images (video, sounds..).
The main limit of this approach is the attempt of validating an inter-subject model, usually BCI calibration and model parameter tuning is performed on the same subject during different session due to the high variability of neural pattern, nevertheless the performance of this model apparently does not exclude different alternative.

References

[1] E. Bijleveld, R. Custers, and H. Aarts, “The unconscious eye opener: Pupil dilation reveals strategic recruitment of resources upon presentation of subliminal reward cues,” *Psychological Science*, vol. 20, no. 11, pp. 1313–1315, 2009.

[2] D. Alnæs, M. H. Sneve, T. Espeseth, T. Endestad, S. H. P. van de Pavert, and B. Laeng, “Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus,” *Journal of vision*, vol. 14, no. 4, pp. 1–1, 2014.

[3] H. Akaike, “A new look at the statistical model identification,” *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716–723, 1974.

[4] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, “The pupil as a measure of emotional arousal and autonomic activation,” *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 2008.

[5] J. Klingner, R. Kumar, and P. Hanrahan, “Measuring the task-evoked pupillary response with a remote eye tracker,” in *Proceedings of the 2008 symposium on Eye tracking research & applications*, 2008, pp. 69–72.

[6] A. A. Zekveld, D. J. Heslenfeld, I. S. Johnsrude, N. J. Versfeld, and S. E. Kramer, “The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load,” *Neuroimage*, vol. 101, pp. 76–86, 2014.

[7] A. Gibaldi, M. Vanegas, P. J. Bex, and G. Maiello, “Evaluation of the tobii eyex eye tracking controller and matlab toolkit for research,” *Behavior research methods*, vol. 49, no. 3, pp. 923–946, 2017.

[8] C. Ellis, “The pupillary light reflex in normal subjects.” *British Journal of Ophthalmology*, vol. 65, no. 11, pp. 754–759, 1981.

[9] R. Dodge and T. S. Cline, “The angle velocity of eye movements.” *Psychological Review*, vol. 8, no. 2, p. 145, 1901.

[10] F. Onorati, R. Barbieri, M. Mauri, V. Russo, and L. Mainardi, “Characterization of affective states by pupillary dynamics and autonomic correlates,” *Frontiers in neuroengineering*, vol. 6, p. 9, 2013.