



Archived at the Flinders Academic Commons:

<http://dspace.flinders.edu.au/dspace/>

This is the authors' version of an article published in *Lecture Notes in Computer Science : Brain Informatics*. The original publication is available by subscription at:

<http://link.springer.com/>

doi : 10.1007/978-3-642-35139-6_20

Please cite this article as:

Atyabi, Adham, Fitzgibbon, Sean, Powers, David, 2012. Multiplication of EEG Samples through Replicating, Biasing, and Overlapping. In *Brain Informatics: International Conference, BI 2012, Macau, China, December 4-7, 2012. Proceedings*. Berlin: Springer Berlin Heidelberg. Pp 209-219

Copyright (2012) Springer Berlin Heidelberg. All rights reserved. **Please note** that any alterations made during the publishing process may not appear in this version.

Multiplication of EEG samples through Replicating, Biasing, and Overlapping

Adham Atyabi¹, Sean P Fitzgibbon¹, and David M W Powers^{1,2}

¹ School of Computer Science, Engineering and Mathematics (CSEM),
Flinders University, Australia.

² Beijing Municipal Lab for Multimedia & Intelligent Software,
Beijing University of Technology, Beijing, China.

{Adham.Atyabi,Sean.Fitzgibbon,David.Powers}@flinders.edu.au

Abstract. EEG recording is a time consuming operation during which the subject is expected to stay still for a long time performing tasks. It is reasonable to expect some fluctuation in the level of focus toward the performed task during the task period. This study is focused on investigating various approaches for emphasizing regions of interest during the task period. Dividing the task period into three segments of beginning, middle and end, is expectable to improve the overall classification performance by changing the concentration of the training samples toward regions in which subject had better concentration toward the performed tasks. This issue is investigated through the use of techniques such as i) replication, ii) biasing, and iii) overlapping. A dataset with 4 motor imagery tasks (BCI Competition III dataset IIIa) is used. The results illustrate the existing variations within the potential of different segments of the task period and the feasibility of techniques that focus the training samples toward such regions.

Keywords: Brain Computer Interface, Replication, Segmentation, Biasing, Overlapping, Triangular Overlapping

1 Introduction

EEG based Brain Computer Interface (BCI) is a communication device that transforms human scalp recordings to executable commands/tasks. EEG recording is a time consuming and tedious procedure during which the subject is expected to sit in a chair and repeatedly perform mental, computational (eg. motor imagery) tasks for some periods of time without being allowed to move. It is unlikely for a subject to consistently maintain a high level of concentration toward the tasks during the task period. That is, it is likely to have some variation in the degree of concentration and task engagement among different time segments of the task period.

Although it is common to assess the performed tasks by evaluating the entire task period as a block epoch, it might be better to only focus toward regions

of the task period during which the subject is expected to have better concentration toward the task. This issue is investigated in [1–3] by proposing various techniques such as segmentation, replication, biasing and overlapping.

In [3] segmentation and replication approaches are applied to a two class motor imagery dataset and it is concluded that the middle and end segments of the task period are more likely to represent higher concentration by subjects toward the task compared with the beginning segments.

In [2] this problem is further investigated through the use of biasing and overlapping techniques. Biasing is reported to be more advantageous compared with replicating since it maintains all sub-epochs originated from the same epoch and only provides higher number of samples within the epoch from regions of interest while in replication approach only samples originated from regions of interest are maintained and the rest are omitted. Similar to [3] the results indicate the potential of middle segments of the task period. In addition, the combination of biasing and overlapping showed better classification performance compared with either of biasing or overlapping. The study was conducted on a dataset with 2 motor imagery tasks.

The conducted studies in [3] and [2] lacked the generalization across datasets and number of classes. This study further evaluate these approaches using a new dataset containing EEG data of 3 subjects performing 4 motor imagery tasks. The outline of the study is as follows. Section 2 introduces the dataset and the applied preprocessing techniques. Section 3 discusses the employed techniques and represents 3 sets of designed experiments and their achieved results. Conclusion and future work are presented in section 4.

2 Dataset

EEG data from the dataset IIIa of BCI Competition III is used [4]. The dataset contains EEG data of 3 healthy subjects (*k3b*, *k6b*, *l1b*) performing a 4 motor imagery task (left hand, right hand, foot, tongue). The EEG dataset is sampled at 250Hz and is band pass filtered between 1Hz and 50Hz. 60 electrodes are used and the task period is set to 3s and 240 tasks are performed in a random order (except with *k3b* that performed 360 trials) [4]. Common average referencing (CAR) and demeaning (D) are the applied preprocessing stages. This is denoted in the paper as CARD250Hz. Frequency features are used and a modified single layer perceptron that incorporates early stopping is employed for classification. All experiments implemented a 10×20 cross validation (CV) which results in 90%, 5%, and 5% ratio for training, validation and testing sets respectively.

Bookmaker informedness is used to assess the classification performance. Detailed discussion about bookmaker can be found in [5–7]. Significance tests (ANOVA if normal or Kruskal-Wallis if not) are performed to investigate the potential of null hypothesis that there is no significant difference among different segments of the epoch.

3 Experimental Design and Results

3.1 Exp. 1: The impact of different sub-windowing

This experiment is designed to investigate the potential of different sub-segments of the task period (epoch) individually. This is implemented by dividing the epoch into non-overlapping sub-epochs with shorter window sizes and the potential of each sub-epoch is assessed individually by being passed to a classifier. Various window sizes are considered (0.8s, 0.6s, 0.5s, 0.4s, 0.3s). The results are demonstrated in Figs 1 and 2. In the figures, *CARD250HzxsRedy*240/360* indicate that the sample rate is 250Hz and common average referencing and demeaning are applied. The sub-epoch length is x second. *Redy* indicate that the illustrated bar represent the y^{th} sub-epoch from the epoch. 240/360 indicate the number of performed tasks (samples). Due to replication in legend test the CRD250Hz and 240/360 are omitted and only the information related to the sub-epoch number is depicted. Therefor, *0.3sRed1* indicate that the epoch is divided to 0.3 s (10 times 0.3s) sub-epochs and among the resulting 10 sub-epochs (each with 0.3s length), the first sub-epoch in each trial is isolated in a new dataset for evaluation. *0.3s All* refers to the use of the entire 10 sub-epochs (0.3s each) in each trial.

The results illustrate variations on individual potential of different segments of the epoch. The use of longer window-sizes resulted in better overall classification performance. This is consistent with previous findings on a 2 motor imagery dataset in [1] and [3]. Using all sub-epochs resulted from dividing the epoch to shorter window sizes showed better classification performance compared to the use of individual sub-epochs. This might be due to the fact that the use of individual sub-epochs reduces the number of training sample compared to the use of all sub-epochs.

3.2 Exp. 2: Replication of regions of interest

The problem of having fewer number of training samples in experiment 1 is resolved in this experiment by replicating regions of interest. To do so, a replicating factor denoted as R is considered (i.e., 10%, 30%, 50%). As the first step, based on the replicating factor, portions of sub-epochs originated from the same epoch are selected from the beginning, the middle, and the end of the epoch forming three new sets. At the next step these sets are replicated to provide new sets with approximately similar number of samples compared with the original set with sub-epochs. The outcome is three new sets that each isolate an specific region of the epoch with approximately equal number of training samples. This procedure is applied to various time windows and the results are illustrated in Fig. 3.

A comparison between Fig. 3 and Figs. 1 and 2 illustrate that the use of higher % as replication factor can improve the classification performance compared to the use of individual sub-epochs. However, in most cases using all sub-epochs out-performed the replicated sets. This might be due to the fact that in the

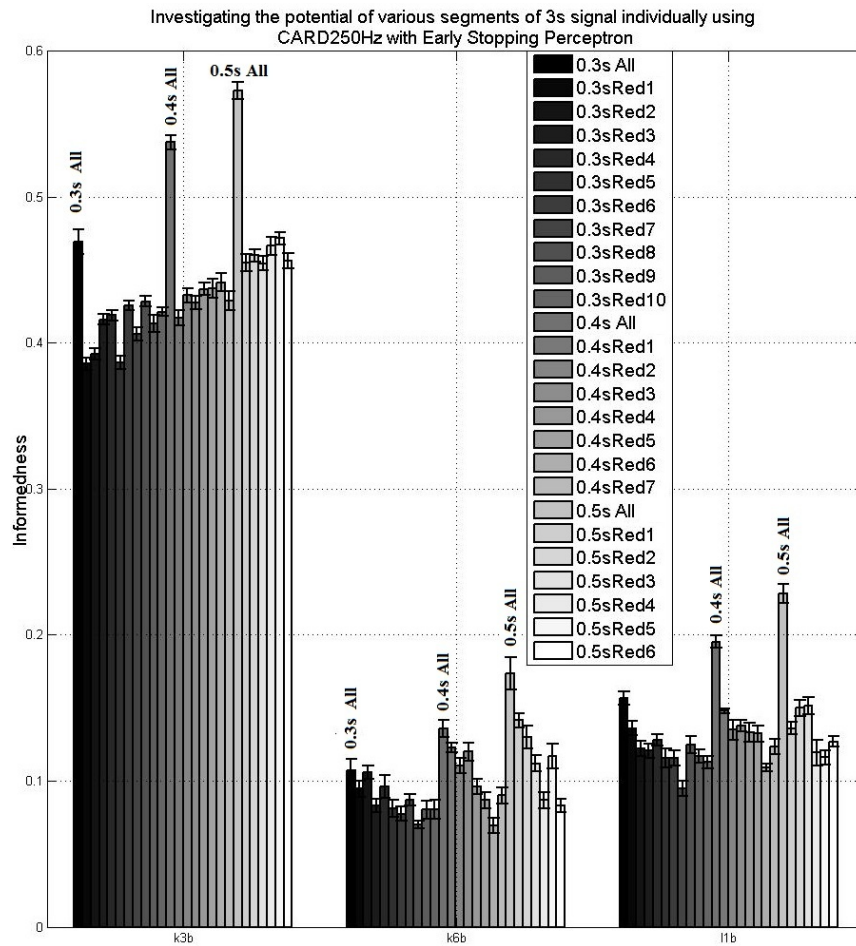


Fig. 1. An investigation on the potential of individual sub-epochs within the epoch using 0.3s, 0.4s, and 0.5s time windows (Exp. 1)

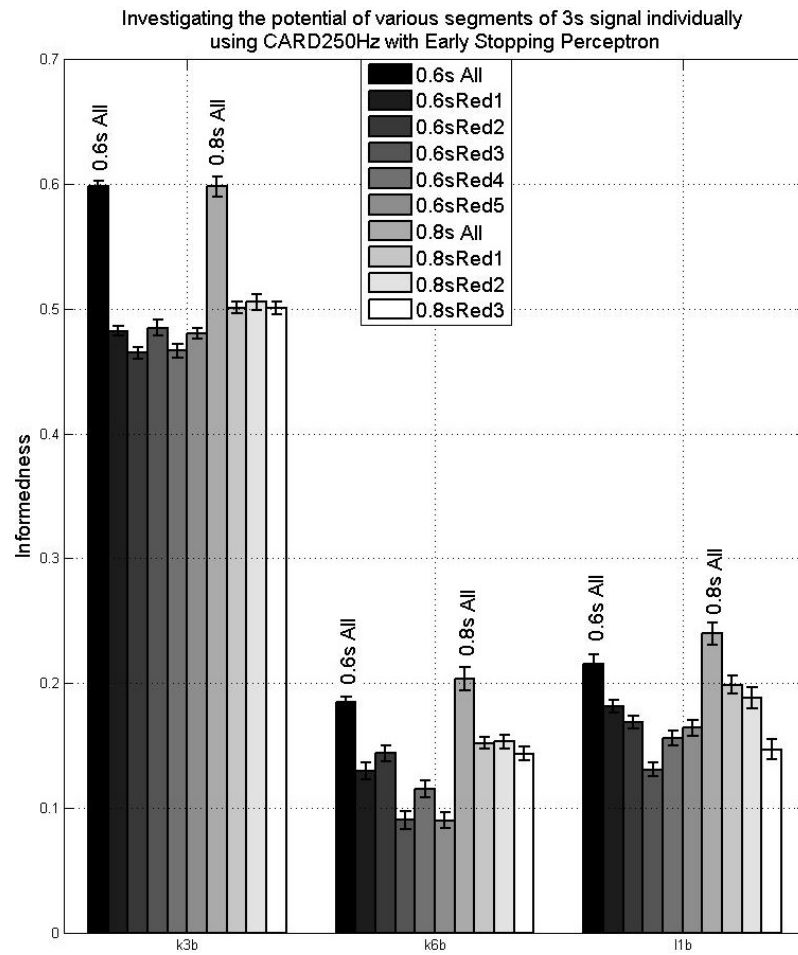


Fig. 2. An investigation on the potential of individual sub-epochs within the epoch using 0.6s and 0.8s time windows (Exp. 1)

replicated sets only sub-epochs from regions of interests are included and the rest are omitted. This issue can be resolved by maintaining all sub-epochs while increasing the number of samples originated from regions of interest (Biasing).

It is noteworthy that unlike the previous findings with 2 motor imagery datasets in [1–3], the beginning and the end segments showed better classification results compared to the middle segments.

3.3 Exp. 3: Biasing toward regions of interest

Exp. 3a: Biased Non-Overlapping vs. non-Biased Overlapping Biasing is a technique that maintains all sub-epochs originated from the same epoch while it provides a higher contribution from those originating from regions of interest in the epoch. The procedure is to divide the epoch into 5 equal segments (B1,B2,M,E2,E1) and implement the biasing in a way that the replication level decreases from regions of interest to the nearby regions. The remaining regions do not get to replicate themselves. Similar to experiment 1, the replication factor is defined as a % (i.e., 10%, 30%, 50%). As an example, using 0.6s sub-windows in 3s epoch generates five sub-epochs within the epoch indexed from 1 to 5. Therefore, B1,B2,M,E2, and E1 are as follows:

$$\left\{ \begin{array}{l} \text{B1}=1^{st} \text{ sub-epoch } (0.0s \dots 0.6s) \\ \text{B1}=2^{nd} \text{ sub-epoch } (0.6s \dots 1.2s) \\ \text{M}=3^{rd} \text{ sub-epoch } (1.2s \dots 1.8s) \\ \text{E2}=4^{th} \text{ sub-epoch } (1.8s \dots 2.4s) \\ \text{E1}=5^{th} \text{ sub-epoch } (2.4s \dots 3.0s) \\ \text{n}=\text{number of sub-epochs within an epoch}=5 \\ \text{p}=\text{round}(\text{percentage} \times \text{n}) \end{array} \right. \quad (1)$$

for $\text{percentage}=50\%$, $p = (50\% \times 5) \approx 3$. The new set would be generated using following criterion:

$$\left\{ \begin{array}{l} \text{Bias toward the beginning} = [2p(B1) \ p(B2) \ p(M) \ E2 \ E1] \\ \text{Bias toward the middle} = [B1 \ p(B2) \ 2p(M) \ p(E2) \ E1] \\ \text{Bias toward the end} = [B1 \ B2 \ p(M) \ p(E2) \ 2p(E1)] \end{array} \right. \quad (2)$$

where $p(x)$ and $2p(x)$ represent p and $2 \times p$ times replicated versions of segment x respectively. In the *CARD250Hz0.6s5*240/360* signal with 3s epoch, the resulting set within each epoch for 50% biasing would be as follow:

$$\left\{ \begin{array}{l} \text{Bias toward the beginning} = [B1 \ B1 \ B1 \ B1 \ B1 \ B1 \ B2 \ B2 \ B2 \ M \ M \ M \ E2 \ E1] \\ \text{Bias toward the middle} = [B1 \ B2 \ B2 \ B2 \ M \ M \ M \ M \ M \ E2 \ E2 \ E2 \ E1] \\ \text{Bias toward the end} = [B1 \ B2 \ M \ M \ M \ E2 \ E2 \ E2 \ E1 \ E1 \ E1 \ E1 \ E1] \end{array} \right. \quad (3)$$

Given that in biasing extra training samples are generated by replicating the regions of interest based on a replication factor while maintaining the remaining sub-epochs (unlike the replication approach), it is possible to assume that any improvement in overall classification performance might be due to having more training samples rather than focusing the training toward regions of interest. To

provide a better assessment, extra results indicating the impact of standard overlapping using various overlapping factors (25%, 50%, 75%, 90%) are included. The results in Fig. 4 indicate that standard overlapping is only superior in higher window sizes (0.8s and 0.6s) while biasing toward regions of interest performs better in smaller window-sizes (0.5s-0.3s). A comparison between Figs. 4 and 3 illustrate the feasibility of biasing compared to replication in terms of achieving a higher classification performance and out performing non-biased, non-replicated sets. The results indicate a significant ($p = 0.0446 < 0.05$) within the potential of various segments of the epoch and biasing factor in addition to existing significant within subjects and time windows ($p < 0.05$).

Exp. 3b: Biased Overlapping windows (Triangular Overlapping) The results achieved in experiment 3a illustrated the potential of both techniques (biasing and overlapping) in terms of providing higher concentration toward regions of interest (biasing) and having higher number of training samples (overlapping). Triangular Overlapping takes advantage of this by biasing regions of interest from the set of overlapped sub-epochs. Similar to previous experiments the biasing is implemented on three regions (the beginning, the middle, and the end). Examples of 25% triangular overlapping toward the beginning and the end on 3s epoch with 1000Hz sample rate using 0.3s window size are illustrated in Figs. 5 and 6.

The comparison between i) non-biased non-overlapped, ii) standard overlapping, and iii) triangular overlapping are provided in Figs. 7 and 8. The results illustrate the potential of triangular overlapping in terms of achieving better classification performance compared to non-biased non-overlapped signal. A comparison between results in Figs. 7, 8, and 4 indicate the superiority of triangular overlapping and no significant difference between 25% triangular overlapping and up to 90% standard overlapping in high window sizes (0.8s) and a better classification performance in lower window sizes. The result of significant test indicate a significant of $p = 0.0001 < 0.05$ within the potential of various percentage of triangular overlapping in addition to existing significant within subjects and time windows ($p = 0 < 0.05$).

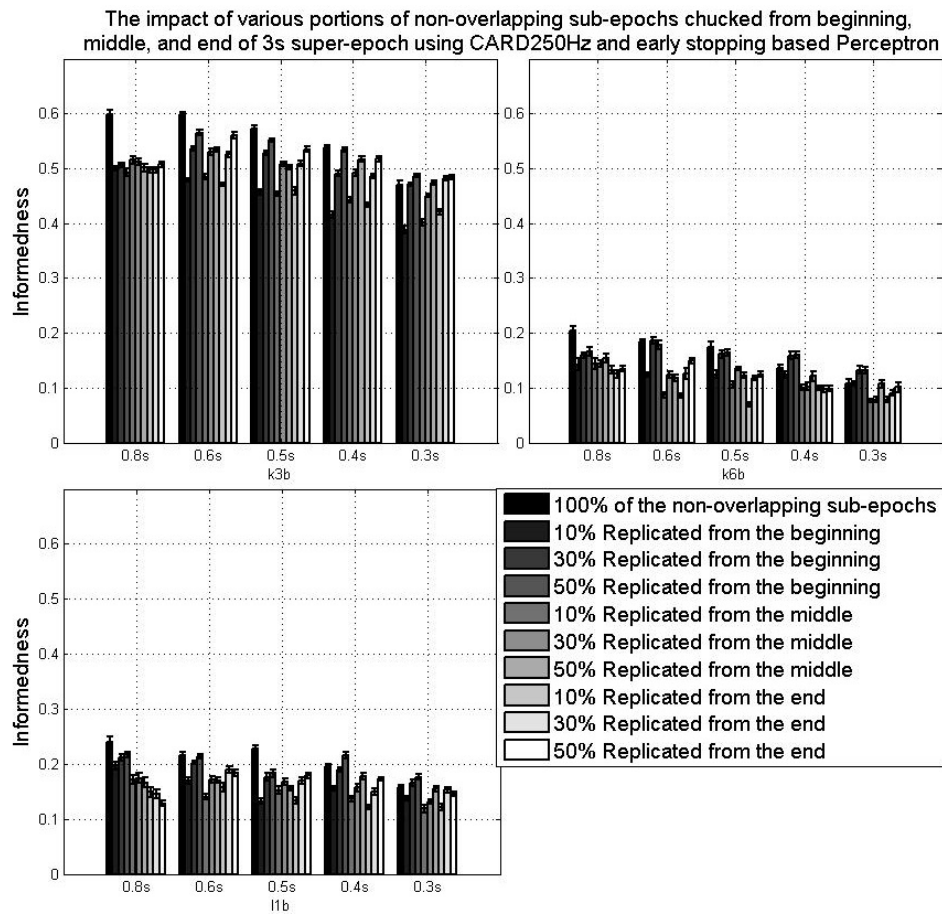


Fig. 3. Comparison of 10%, 30%, and 50% replication from beginning, middle and end segments of epoch and the use of the entire sub-epochs on various window sizes (0.8s, 0.6s, 0.5s, 0.4s, 0.3s) (Exp. 2)

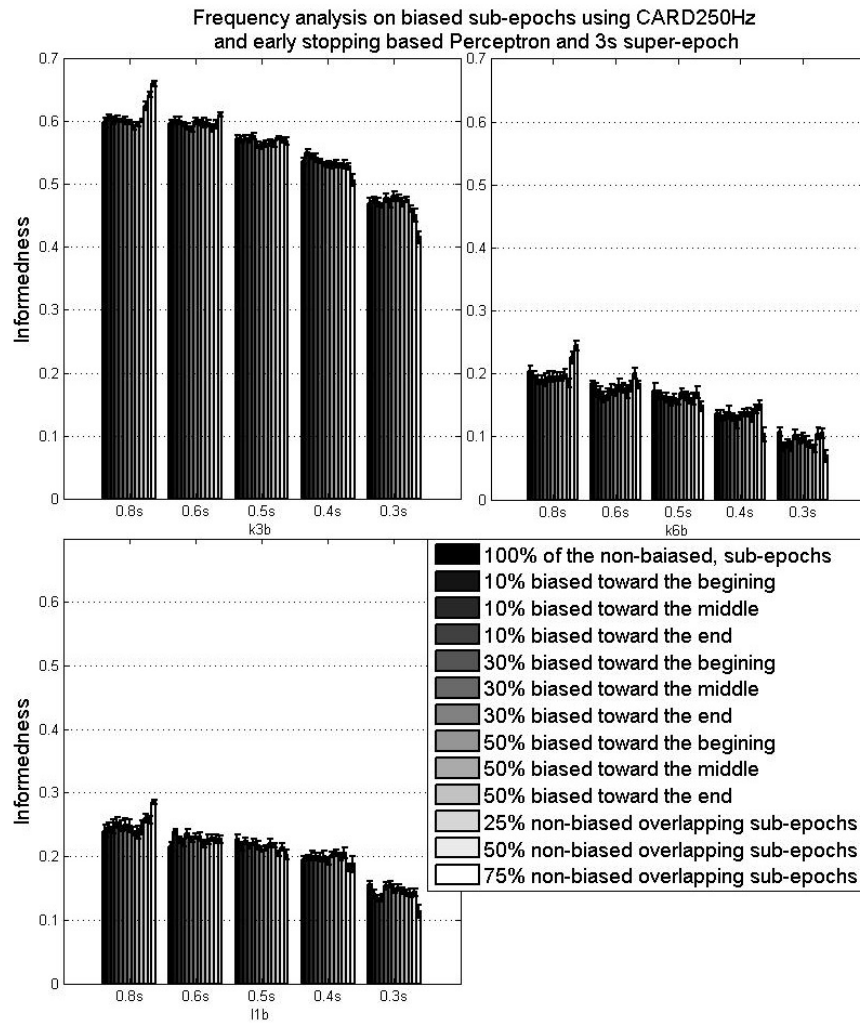


Fig. 4. 10%, 30%, and 50% biasing toward beginning, middle and end segments of epoch vs. 25%, 50%, and 75% standard overlapping using various window sizes (0.8s, 0.6s, 0.5s, 0.4s, 0.3s) (Exp. 3a).

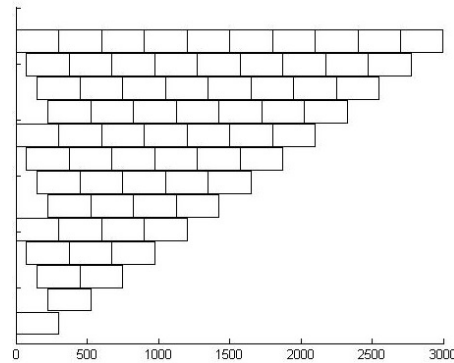


Fig. 5. Triangular Overlapping toward the beginning resulting in 67 sub-epochs using a 0.3s window size on 3s epoch with 1000Hz sample rate as an example (Exp. 3b).

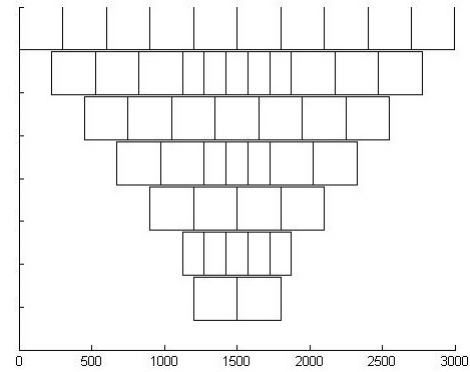


Fig. 6. Triangular Overlapping toward the middle resulting in 48 sub-epochs using a 0.3s window size on 3s epoch with 1000Hz sample rate as an example (Exp. 3b).

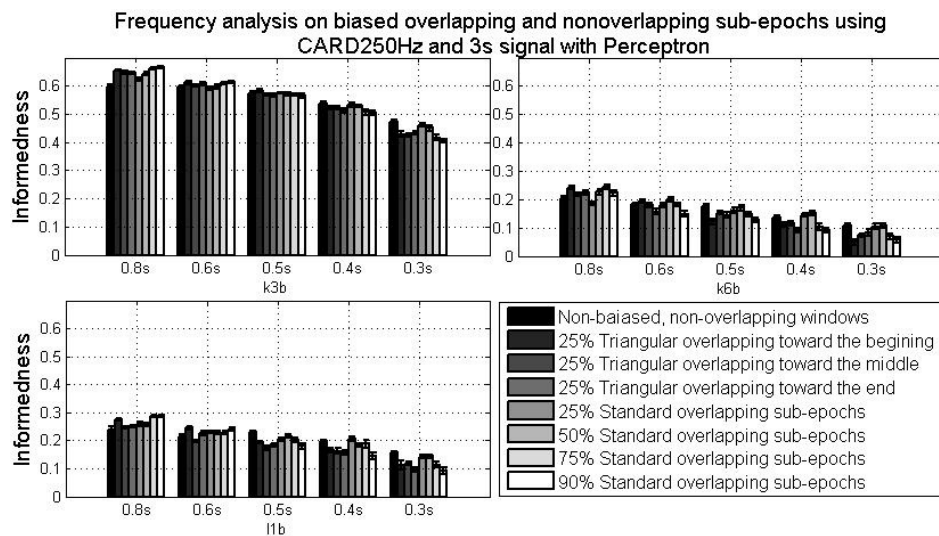


Fig. 7. A comparison of 25% triangular overlapping toward beginning, middle and end segments with various % of standard overlapping using various window sizes (0.8s ,0.6s, 0.5s, 0.4s, 0.3s) (Exp. 3b).

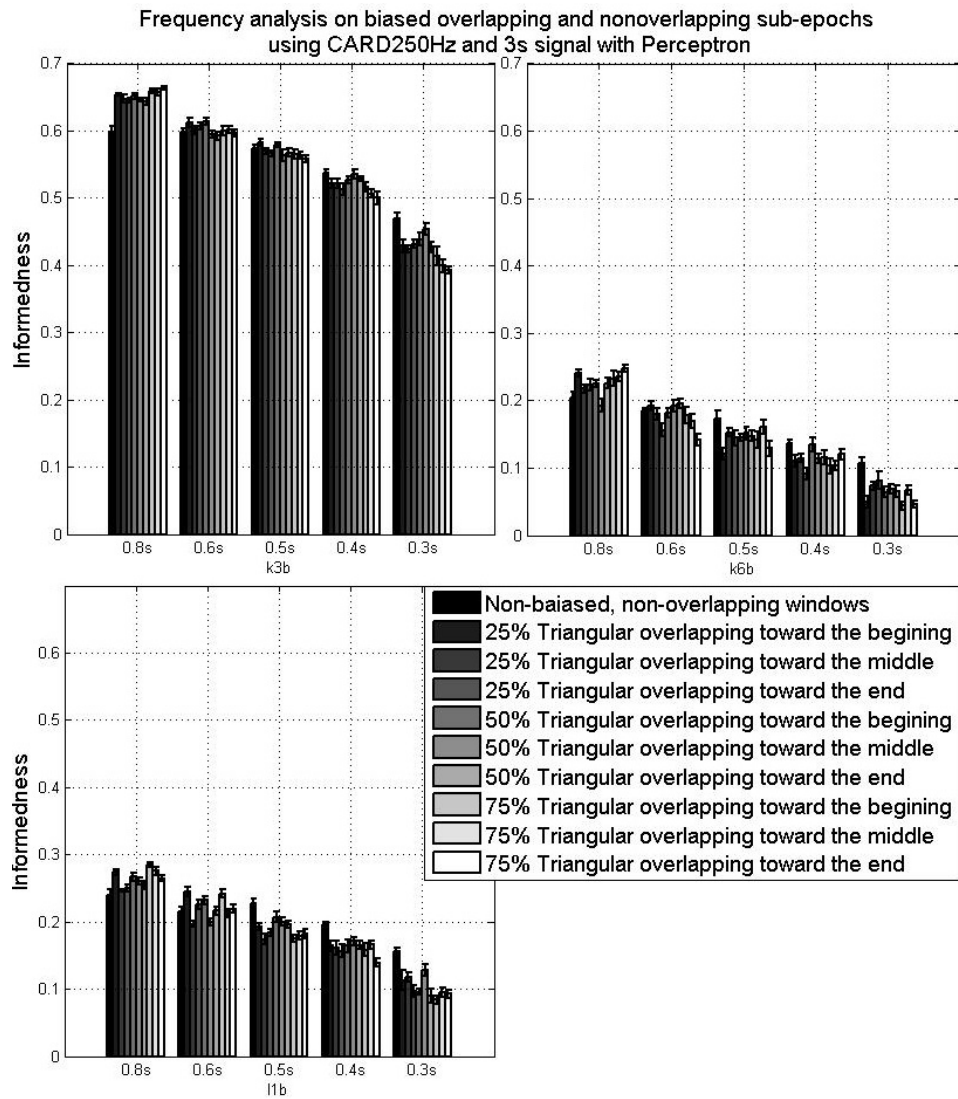


Fig. 8. A comparison of various % of triangular overlapping toward beginning, middle and end segments using various window sizes (0.8s ,0.6s, 0.5s, 0.4s, 0.3s) (Exp. 3b).

4 Conclusion

This paper investigated the potential of several approaches to change the concentration of the training set toward regions of interest in which it is expected from subjects to have higher concentration on the tasks. Replicating, Biasing, and Triangular Overlapping are examined using a dataset of 3 subjects performing 4 motor imagery tasks. The conducted experiments illustrated variations within different segments of the epoch (beginning, middle, and end). Among the examined approaches triangular overlapping achieved better classification performance. This is likely to be due to providing higher concentration toward regions of interests in the epoch (caused by biasing) in addition to increasing the number of training samples (caused by overlapping). The variation in achieved performance by different segments of the epoch is inconsistent with the previous findings with these approaches on a dataset of 5 subjects performing 2 motor imagery tasks in [1–3] and indicate the higher potential of the beginning and the end segments in most cases.

References

1. Atyabi, A., Fitzgibbon, S. P., Powers, D. M. W.: Multiplying the milage of your dataset with subwindowing, Proceedings of the 2011 international conference on Brain informatics (BI'11), Springer Lecture Notes in Computer Science, Vol. 6889, pp. 173–184, (2011)
2. Atyabi, A., Fitzgibbon, S. P., Powers, D. M. W.: Biasing the Overlapping and Non-Overlapping Sub-Windows of EEG recording, In: IEEE International Joint Conference on Neural Networks (IJCNN'12), The 2012 IEEE World Congress on Computational Intelligence (IEEE WCCI 2012), pp. 741–746, (2012)
3. Atyabi, A., Powers, D. M. W.: The impact of Segmentation and Replication on Non-Overlapping windows: An EEG study, The Second International Conference on Information Science and Technology (ICIST2012), China, pp. 668–674, (2012)
4. Blankertz, B., Muller, K.-R., Krusienski, D. J., Schalk, G., Wolpaw, J.R., Schlogl, A., Pfurtscheller, G., del R. Millan, J., Schroder, M., Birbaumer, N.: The BCI competition III: Validating alternative approaches to actual BCI problems. *Neural Syst. Rehabil. Eng.*, **14**(2), pp. 153–159, (2006)
5. Powers, D. M. W.: Recall and Precision versus the Bookmaker. *International Conference on Cognitive Science (ICSC-2003)*, pp. 529–534, (2003)
6. Powers, D. M. W.: Evaluation: From Precision, Recall and F-Measure to ROC., Informedness, Markedness & Correlation, *Journal of Machine Learning Technologies*, **2**(1), pp. 37–63 (2011)
7. Powers, D. M. W.: The Problem of Kappa, 13th Conference of the European Chapter of the Association for Computational Linguistics, Avignon France, April (2012)