Target Detection in Video Feeds with Selected Dyads and Groups Assisted by Collaborative Brain-Computer Interfaces

Saugat Bhattacharyya, Davide Valeriani, Caterina Cinel, Luca Citi and Riccardo Poli

Abstract—We present a collaborative Brain-Computer Interface (cBCI) to aid group decision-making based on realistic video feeds. The cBCI combines neural features extracted from EEG and response times to estimate the decision confidence of users. Confidence estimates are used to weigh individual responses and obtain group decisions. Results obtained with 10 participants indicate that cBCI groups are significantly more accurate than equally-sized groups using standard majority. Also, selecting dyads on the basis of the average performance of their members and then assisting them with our cBCI halves the error rates with respect to majority-based performance. Also, this allows most participants to be included in at least one selected dyad, hence being quite inclusive. Results indicate that this selection strategy makes cBCIs even more effective as methods for human augmentation in realistic scenarios.

I. INTRODUCTION

Research on collective intelligence has shown that groups are generally superior to individuals in a variety of decision-making tasks involving animals and humans [?]. These include estimating uncertain quantities [?], perceptual tasks [?], [?], and medical decision making [?].

To exploit their augmented capabilities and intelligence, groups should either be allowed to communicate freely [?] or have individuals decide in isolation and make collective judgements taking into account the decision confidence of each group's member [?]. However, communication between group members is not always possible, and calibrated decision confidence estimates (i.e., unbiased predictors of decision correctness [?]) may not be available [?], [?]. In these cases, group decision making could be suboptimal [?].

Recently we have shown that it is possible to use collaborative Brain-Computer Interfaces (cBCIs) to estimate the decision confidence from the brain signals and response times (RTs) of isolated users performing a matching task [?]. We used confidence estimates to weigh individual decisions and make group decisions, which were significantly better than those of individuals and equally-sized groups using standard majority. Similar results were then obtained with more realistic tasks based on static images [?], [?], [?]. We also showed that cBCI confidence estimates were better calibrated than the confidence reported by participants after each decision [?].

The authors are with the Brain Computer Interfaces and Neural Engineering Laboratory, School of Computer Science and Electronic Engineering, University of Essex, UK. Email: {saugat.bhattacharyya,dvaler,ccinel,lciti,rpoli}@essex.ac.uk.

While these previous studies focused on the performance of *average* groups assisted by cBCIs without any form of selection, it is clear that certain groups perform better than others. For example, dyads (i.e., pairs of users) or triads made by individuals with similar performance were found to exhibit maximum collective benefits in [?], [?].

This study extends previous work on cBCIs for augmented decision making along two directions. Firstly, we investigate whether cBCIs could improve group performance when users are exposed to *video feeds* from a realistic, dynamic environment. Secondly, we introduce a method for *selecting groups* that maximises performance without excluding *a priori* any individual.

II. METHODOLOGY

A. Participants

Twelve healthy participants (seven females, four left-handed, age = 35.4 ± 7.4 years), with normal or corrected-to-normal vision and no reported history of epilepsy, took part in a decision-making experiment. All volunteers signed an informed consent form before taking part in the study and were paid GBP 16 for approximately 2 hours of their time. Data from two participants were discarded because the EEG signals were too noisy or did not get recorded at all. This experiment received Ministry of Defence Research Ethics Committee (MODREC) ethical approval in July 2017.

B. Stimuli and Procedure

Participants were presented with 12 blocks of 28 trials, for a total of 336 trials. In each block (Fig. 1), a video sequence was presented (with a 4 Hz frame rate) representing the viewpoint of a user walking along a corridor, where individuals can appear from doorways located at both sides. When a person appeared, participants had to decide, within 2.5 s, whether the individual was wearing a helmet or a cap by pressing the left or the right mouse button, respectively. After the decision, they were asked to indicate, within 2 s, the degree of confidence in their decision (0-100%) using the mouse wheel (although these data were not used in this study). The video sequence presented in each block contained 42 pairs of doorways: one third with individuals wearing a helmet, one third with individuals wearing a cap, and one third empty (no decisions required), in random order.

Participants were comfortably seated at about 80 cm from an LCD screen. Prior to the start of the experimental session, each participant underwent a brief training session of 14 trials to familiarise with the task.
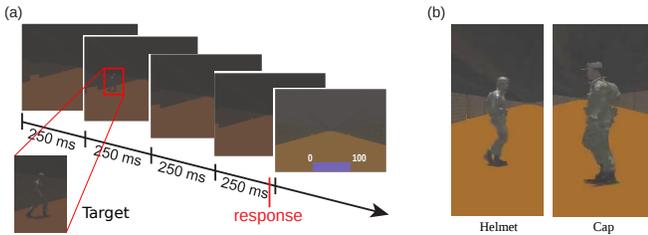
Fig. 1: (a) Sequence of stimuli presented in a trial. (b) Examples of the avatars wearing a helmet (left) and a cap (right) used in the experiment.

C. Data Acquisition and Processing

A Biosemi ActiveTwo EEG system was used to record the neural signals from 64 electrode sites following the 10-20 international system. The EEG data were sampled at 2048 Hz, referenced to the mean of the electrodes placed on the earlobes, and band-pass filtered between 0.15 and 40 Hz. Artefacts caused by ocular movements were removed using a standard subtraction algorithm based on correlations. Then, response-locked epochs starting 1 s before the response and lasting 1.5 s were extracted from the neural data associated to each trial, detrended, baseline corrected using the average voltage recorded in a time window of 200 ms before the beginning of the trial, and finally filtered and downsampled to 32 Hz. Each epoch was then labelled as *confident* or *not confident*, depending on whether the participant made a correct or incorrect decision, respectively.

Common Spatial Pattern (CSP) [**?**] was used to extract neural features from each epoch. We used 8-fold cross validation to split the data into training and test sets, to ensure the results were not affected by over-fitting. The training set is used to compute the CSP matrix, which transforms the response-locked EEG data into a lower-dimensional subspace where the two classes (correct and incorrect) are more distinguishable. The logarithm of the variances of the first and the last rows of the CSP-transformed data have then been used as neural features in our cBCI to predict the confidence of the user in each decision. The variances for the two classes (i.e., correct and incorrect responses) are largest in the first and the last rows of the subspace and, hence, they are most distinguishable. Thus, the logarithm of the variances of the first and the last spatial subspaces were used as neural features in our cBCI.

We studied the performance of both a cBCI where the feature vector was composed by these two CSPs only, and a cBCI using the RT of the participant, which is known to correlate with the decision confidence [**?**], in addition to the two CSPs. RTs were measured by timestamping the click of an ordinary USB mouse.

D. Confidence Estimation and Group Decisions

For each participant p, a logistic regression model was used to predict the confidence weight of trial i, $w_{p,i}$, from the feature vector. The model was fitted using L2 normalization and a regularization strength $C{=}1000$. Group decisions were then made as follows:

$$d_{group,i} = \mathtt{sign}\left(\sum_{p=1}^{m} w_{p,i} \cdot d_{p,i}\right),$$

where $d_{p,i}$ is the decision of participant p in trial i, $w_{p,i}$ is the corresponding confidence weight.

Groups of size $m{=}2,\ldots,10$ were formed offline by considering the $\binom{10}{m}$ combinations of the 10 participants. The performance of groups assisted by the cBCI based only on CSP and the one also based on RT was then compared with the performance obtained by traditional groups using standard majority (i.e., $w_{p,i}{=}1$, for all p and i).

E. Dyad Selection

In the case of dyads (i.e., groups of size 2), we also studied the performance of the subset of the original 45 dyads which had standard-majority performance, β, over a certain threshold. We should note that the majority performance of dyads is the average performance of their members.

When a dyad is asked to make a decision in a binary decision-making task, three possible scenarios can happen:

1) both users make the correct decision;
2) both users make the incorrect decision;
3) users make opposite decisions, hence generating a *tie*.

A tie represents an opportunity for the dyad to correct a mistake made by one of the members if an effective tie-breaking strategy is available to decide which member should be trusted. Standard majority uses coin-flipping to break ties, hence turning only half of the ties into correct decisions. Its performance is, therefore, given by:

$$\beta = S_{11} + \frac{1}{2} \cdot T,$$

where S_{11} is the fraction of decisions in which both members made a correct decision and T is the fraction of ties. Dyads with high β have a high number of unanimous correct decisions S_{11}, a high percentage of ties T that were turned into correct decisions by a good tie-breaking strategy (such as our cBCIs), or a combination of both.

Our proposed selection method for cBCI-assisted dyads is to accept only dyads with $\beta > \beta_t$, where β_t is a constant. While, by definition, it is clear that the higher the β, the better the performance of a dyad *when using majority*, it was not immediately obvious whether such a selection strategy would also lead to good cBCI-assisted groups.

III. RESULTS

A. Individual Performance

Fig. 2 shows the error rates of 10 participants in the experiment. On average, the participants made incorrect decisions in 21.55% of the trials. Seven of the ten participants had error rates below the mean, the mean being negatively affected by participant 3, 5 and 6 who presented error rates of 30% or more.
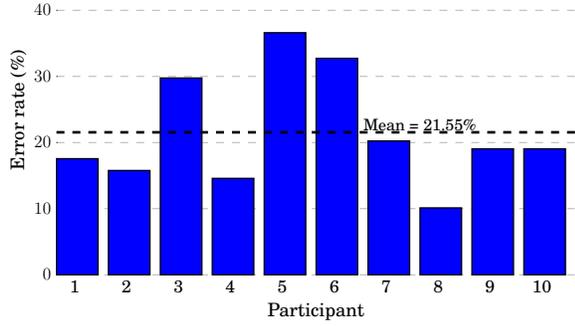
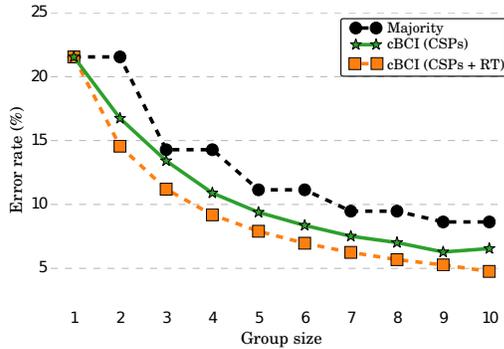Fig. 2: Error rates for each participant of the experiment.

Fig. 3: Error rates obtained by groups of sizes 1 to 10 when using standard majority (black), a cBCI based on CSP features (green), and a cBCI based on CSP and RT features (orange).

B. Group Performance

Fig. 3 shows the error rates of groups of size 1 to 10. The difference between error distributions within each group size for different methods were statistically tested using the one-tailed Wilcoxon signed-rank test.

The results indicate that, for groups of sizes 2 to 9, the cBCI based on neural features (labelled "cBCI (CSPs)" in the figure) is significantly better than the standard majority method (Wilcoxon $p<0.003$). Moreover, the cBCI based on both neural features and RT (labelled "cBCI (CSPs+RT)") outperformed the cBCI based solely on the neural features ($p<0.004$) for groups of sizes 2 to 9.

C. Performance of Selected Dyads

Fig. 3 also showed that dyads assisted by the cBCI based on neural features and RT reduce their error rates from 21.55% (obtained using standard majority) to 14.56% thanks to the cBCI breaking ties better than random. This is the performance improvement obtained without any selection.

Fig. 4 compares the error rates of selected dyads using standard majority and the cBCI based on neural features and RT, for different threshold values β_t. The number of dyads selected when adopting each threshold values are shown at the top of the figure. The error rates of cBCI-assisted dyads were significantly lower than those obtained
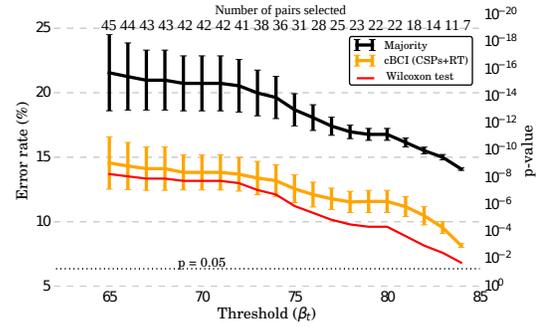

Fig. 4: Error rates of selected dyads using standard majority (black) and assisted by the cBCI based on neural features and RT (orange) for increasing values of the threshold β_t. The corresponding p-values of the Wilcoxon signed-rank test comparing the two methods are also shown in red. Sample sizes are shown at the top.
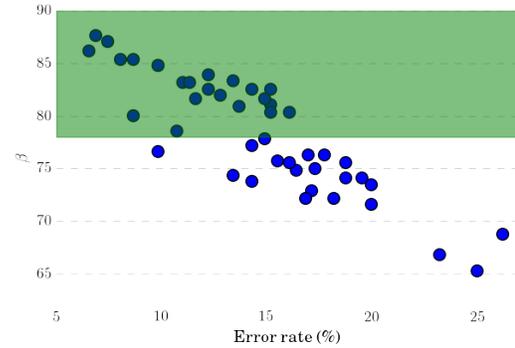

Fig. 5: Majority performance of all 45 dyads as a function of the error rates obtained when assisting dyads with the cBCI based on neural features and RT. The green region marks the dyads selected when the threshold β_t is set to the average individual performance.

with standard majority for all threshold values (one-tailed Wilcoxon signed-rank $p < 0.05$, red line in the figure). Interestingly, error rates decrease as β_t increases also when considering cBCI-assisted dyads.

To further analyse the effects of our group selection strategy, Fig. 5 reports the values of β of all original dyads as a function of the error rates obtained using the cBCI based on neural features and RT. The green area shows which dyads are selected when setting $\beta_t = 78.45\%$ (average individual performance). The results clearly show that the best cBCI-assisted dyads (left-most dots, i.e., low error rates) are also the best majority-based dyads (top-most dots, i.e., high β). Hence by selecting dyads using our strategy based on behavioural performance, one also approximately selects the optimal cBCI-assisted dyads.

D. Membership of Selected Dyads

Fig. 6 shows the number of dyads in which each participant was a member after performing dyad selection with $\beta_t=78.45\%$. Without selection, each participant took part in nine dyads. After selection, all users except for participant 5
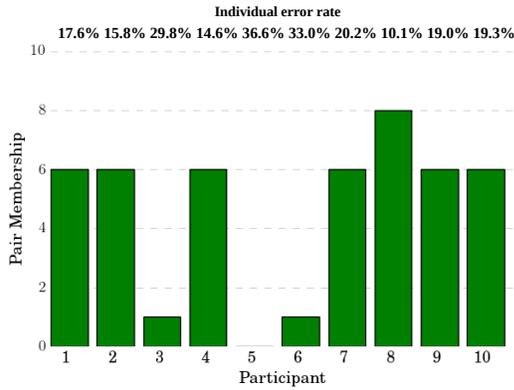
Individual error rate

17.6% 15.8% 29.8% 14.6% 36.6% 33.0% 20.2% 10.1% 19.0% 19.3%

Fig. 6: Number of dyads in which each participant was a member after selection using $\beta_t{=}78.45\%$. The error rates of each participant are indicated on top.

still appeared in at least one dyad. The reason why participant 5 was excluded by our strategy is that it was the user with the highest individual error rate. Hence, dyads including this participant were more likely to perform worse than the average individual when using standard majority. The second- and third-worst performers (i.e., participants 3 and 6) worked well only with the best performer (participant 8), who worked well with all other participants (except 5). The

In future research, we plan to verify that our group selection method also works in other decision-making tasks, and to extend it to work with bigger groups. Furthermore, other methods for group selection could be introduced, e.g., only selecting those participants that exhibit robust brain patterns correlated with the decision confidence. This could

other participants worked well with one another. We should note that had we selected *individuals*, instead of *dyads*, based on whether their performance was above average, we would have excluded participants 3, 5 and 6 completely.

IV. CONCLUSION

This paper has showed that weighing individual decisions by the confidence estimated by a collaborative brain-computer interface using neural features and response times can significantly improve group performance even in realistic decision environments involving video feeds. The biggest improvement brought by the cBCI over majority occurs in dyads thanks to the tie-breaking strategy yielded by the cBCI being better than random. These findings make another step towards applying cBCIs in real-world critical tasks, such as threat detection.

The performance of cBCI-assisted dyads was further improved by selecting only the dyads with a majority performance above a certain threshold. The error rate of such dyads when the threshold was set to the average individual accuracy and when they were assisted by our cBCI was nearly halved w.r.t. the error rate of dyad using standard majority and in the absence of selection. Also, while our selection strategy discarded half of the original dyads, it only discarded one participant, hence being very inclusive.

lead to methods for "neural fingerprinting", which could be combined with our group selection strategy to further enhance group decision making accuracy.

REFERENCES