Experiences and Insights from the Collection of a Novel Multimedia EEG Dataset

Graham Healy, Zhengwei Wang, Tomas Ward, Alan Smeaton, and Cathal Gurrin

Insight Centre for Data Analytics, Dublin City University, Glasnevin, Dublin, Ireland
{graham.healy[**],zhengwei.wang,tomas.ward,alan.smeaton,cathal.gurrin}@dcu.ie

Abstract. There is a growing interest in utilising novel signal sources such as EEG (Electroencephalography) in multimedia research. When using such signals, subtle limitations are often not readily apparent without significant domain expertise. Multimedia research outputs incorporating EEG signals can fail to be replicated when only minor modifications have been made to an experiment or seemingly unimportant (or unstated) details are changed. This can lead to overoptimistic or overpessimistic viewpoints on the potential real-world utility of these signals in multimedia research activities. This paper describes an EEG/MM dataset and presents a summary of distilled experiences and knowledge gained during the preparation (and utilisiation) of the dataset that supported a collaborative neural-image labelling benchmarking task. The goal of this task was to collaboratively identify machine learning approaches that would support the use of EEG signals in areas such as image labelling and multimedia modeling or retrieval. The contributions of this paper can be listed thus; a template experimental paradigm is proposed (along with datasets and a baseline system) upon which researchers can explore multimedia image labelling using a brain-computer interface, learnings regarding commonly encountered issues (and useful signals) when conducting research that utilises EEG in multimedia contexts are provided, and finally insights are shared on how an EEG dataset was used to support a collaborative neural-image labelling benchmarking task and the valuable experiences gained.

Keywords: Brain-computer Interface · Electroencephalography · RSVP

1   Introduction

EEG (Electroencephalography) has recently become an accessible method to support the building and operation of BCI (Brain-Computer Interface) applications. While the initial use of such techniques began in clinical / rehabilitative settings for the purposes of augmenting communication and control, a recent trend has been to use such signals and methods in novel domains, such as image annotation, which relies on the identification of target brain events to trigger

---

[**] Please contact Dr. Graham Healy for enquires related to this paper.

semi-automated image labeling [5][11][14][19]. This trend is particularly relevant to multimedia and HII (Human-Information Interaction) communities because in recent years EEG has demonstrated its potential for several applications including annotation of multimedia content, identifying when a user's attention is drawn to something in the real world, or as a source of wearable sensor data to be indexed for later retrieval or analysis.

EEG signals when used in a multimedia application can often be naively expected to carry meaningful information that will directly measure a particular mental state or concept. Many multimedia researchers embark on a course of research intending to use EEG signals without a clear understanding of which phenomena in the signal should be useful. In such circumstances, it is often quickly realised that the signals offer little utility. Conversely, a naively applied off-the-shelf machine-learning strategy to an arbitrary selection of features from EEG might not readily reveal if the source of the useful information is of neural origin or from non-neural artefacts imparted onto the EEG (e.g. eye movements), therefore it can be erroneously assumed that since EEG was used the only explanation is that the useful signals are directly of a neural origin.

EEG data is rife with sources of variability including those related to the task, the environment and the participants themselves. Moreover, EEG is typically contaminated by non-neural sources of activity emerging from the body such as eye movements (EOG - Electrooculography), facial movement (EMG - Electromyography) and ECG (Electrocardiography). These signals are typically orders of magnitude larger than EEG phenomena, and can be difficult to disentangle from the EEG. Hence, we decided that the community needed to have access to multimedia datasets containing EEG signals and a better understanding of the methodologies for extracting semantic value from such data. Therefore, an EEG/multimedia modelling comparative benchmarking task was proposed and run at NTCIR-13.

NTCIR (NII Testbeds and Community for Information access Research) is a repeating participation conference (with an 18-month cycle) that brings together researchers to develop evaluation methodologies and performance measures for IA (Information Access) technologies across single and multiple-media data. This brings together an active research community in which findings based on comparable experimental results are shared and exchanged in an open, collaborative manner. One topical focus of this is mining knowledge from a large amount of human generated data. NAILS (Neurally Augmented Image Labelling Strategies) was an affiliated task at NTCIR-13 (the 13th NTCIR conference) to support the collaborative evaluation of best-practice strategies for RSVP-EEG image search applications, where researchers benchmarked their machine-learning strategies.

This paper extends previous work with the NAILS task by sharing experiences gained in collecting the novel dataset, and importantly, offering not only a template experimental protocol but also an understanding of common pitfalls and issues. We describe the experimental protocol used to capture the dataset for this task, and compliment this with an understanding of the motivation be-

hind its construction to allow others to extend upon this approach integrating it as relevant with their application domain.

2   Motivation to use EEG

When coupled with a suitable visual presentation paradigm, EEG can enable the detection of attention-related events that are understood to be indicative of user interest – or more specifically the allocation of their attention to one particular stimulus as opposed to some other. One characteristic pattern of activity, commonly known as the P300 signal [20], has been a focus of investigation as it can be used as an index of attentional resource allocation to a stimulus such as an attention-captivating image (due to its infrequency) when presented on a screen. This finding has enabled BCI systems to leverage the ability of a user to be able to guide their attention in such a manner so as to be able to provide relevance judgements/ratings on visual stimuli. For example, a user can actively 'look out' for a particular type of image so that when relevant images appear in a high-speed visual presentation sequence known as RSVP (Rapid Serial Visual Presentation) [26], they will subsequently elicit a P300 response that can be detected using signal processing and machine-learning methods. Ultimately this allows the image to be 'neurally' labelled by the participant.

While systems like these have been explored in a proof-of-concept manner in BCI research using a multitude of image-search tasks, the datasets used usually remain unshared between studies, making it difficult to meaningfully compare the machine-learning and feature-processing strategies used, to find those that offer the best generalisability both across tasks and participants. This is what NTCIR-13 NAILS sought to address, by developing and releasing a dataset and setting achievable research challenges for participants. While the NAILS task showed some classification approaches from participating groups were better than others, it also provided clear evidence, that even with a pristine cleaned dataset (removing potential non-neural sources of useful information), that applications relying on such signals would still need to ultimately rely on noisy labels for the time being i.e. a balanced accuracy of 0.8839 might not be good enough yet to support the type of image target detection applications it is targeting.

While this might appear as a roadblock for applications using EEG signals in this way, it's important to realise the developments that have been made in this space over the past number of years, where through refined experimental protocols, better sensing equipment [15][16][10] and improved signal processing/machine-learning [12], the accuracy of such neural-labelling systems are improving. Moreover, applications replying upon the RSVP-EEG paradigm are being realised where the objective of the system might be to extract other information such as the believability of images generated by GANs [28], memorability of media [24], or to use the signals across users on the same dataset collectively to overcome issues with noise [8]. It is the opinion of the authors that sharing datasets, as well as experience and knowledge in how to conduct EEG

data gathering and use of the data from a RSVP-EEG BCI, could enable other researchers to leverage a RSVP-EEG BCI as a component to drive other applications within the multimedia modelling space although the intended application may not be specifically for image labelling.

Fig. 1. Examples of four target images used in the NAILS experiment.

3   NAILS Data Set & Collection

Before discussing experiences in executing the NAILS task at NTCIR-13, we provide a concise description of the dataset [6] [7].

3.1   Experimental Task Description

The NAILS dataset[1] collection contained EEG responses to 97,200 images from 10 experimental participants. Data collection was carried out with approval from Dublin City University's Research Ethics Committee (DCUREC/2016/099). Each participant completed 6 different search tasks for a particular type of target,

---

[1] To gain access to the NAILS dataset and related benchmark implementations please contact graham.healy@dcu.ie

where each search task was divided into 9 (approximately 35 second) blocks which were completed in a self-paced manner so as to alleviate strain on participants. In each search task, a participant searched for a known type of target (e.g. an airplane), and was instructed to covertly count occurrences of target images in the RSVP sequence so as to maintain their attention on the task. Figure 1 shows examples of the target search images used. In each RSVP block, images were presented successively at a rate of 6 Hz with target (search-relevant) images randomly interspersed among standard (non-search relevant) images with a percentage of 5% across all blocks. In each block, 180 images (9 targets/171 standards) were presented in rapid succession on screen. Per participant, there were 486/9234 target/standard examples available.

3.2   Image Dataset

Image tasks were constructed using freely available datasets [29][23][22]. These were selected as a good choice given that they are commonly used datasets with well-researched characteristics that are representative of the visual content typically encountered in multimedia-IR tasks whilst remaining similar to content used in previous RSVP-BCI studies.

3.3   EEG data filtering

As contaminant eye-movement related activity on the EEG can often contain useful information, epochs (from -1000ms to 2000ms) containing such activity were excluded as they might encourage developed strategies to utilise these non-neural sources of discriminative information. In the NTCIR-13 NAILS task, epochs were filtered to exclude those with a peak-to-peak amplitude greater than 70 on EOG (Electrooculogram) and frontal EEG channels to remove trials that contained such contaminant eye-movement activity. A commonly employed strategy in EEG signal processing is ICA (Independent Component Analysis), and in the NAILS dataset this was used alongside a wavelet based analysis to confirm that the remaining epochs did not contain non-neural sources of discriminative information. For further details please refer to [7]. For the NAILS task, this dataset was split into a training/testing set, where 15/285 target/standard trials from each search task (for each participant) were selected to act as a withheld test set in the evaluation.

3.4   Collaborative Evaluation Task Description

Nine competing teams took part in the collaborative evaluation using the supplied training data (remaining epochs from blocks not used to extract test set data) and they were asked to build machine-learning models that maximised the BA (Balanced Accuracy) score on the withheld testing set (withheld by the NAILS organisers). That means for an evaluation run, a team needed to submit binary predictions for the 18,000 examples given in the test set (900/17100 targets/standards respectively). There were more than 2500/47000 target/standard training examples available across all participants for model training.

3.5   Provided Features / Pre-processing

Three types of preprocessed data were made available to participating organisations: time-series features (time), wavelet magnitude features (w-mean) and wavelet magnitude ratio (w-ratio) features. It was at the discretion of each participating team which combination of these features to use.

3.6   Dataset Validation

In order to validate that the captured data contained useful information for classification prior to sharing the dataset, we applied a basic machine learning analysis using a RBF (Radial Basis Function) kernel SVM (Support Vector Machine) [18]. Each model was trained on a participant-by-participant basis where hyper-parameters (C and gamma) were learned using a randomised grid-search approach. Each model was then applied to the unseen test set data where accuracy measures were calculated (presented in Table 1 and Table 2). A range (and combination) of feature sources (used in baseline approaches) were presented so as to support a better interpretation of the results of participating teams. Importantly, a functioning pipeline was shared with participants (developed in python using mne [3] and sklearn [17]) demonstrating how this result was achieved. In Figure 2, we show a characteristic P3b response acquired from one experimental participant. These measures verified that the chosen tasks were eliciting the expected characteristic (oddball) P300 response i.e. it was possible for a participant to do the search tasks as intended.

Table 1. Balanced accuracy scores for each participating team's best performing method broken down by experimental participant.

Dataset	Team-1	Team-2	Baseline
101	.8219	.7670	.7503
102	.8781	.8512	.8211
103	.8646	.8275	.7664
104	.8877	.8743	.8322
105	.9304	.8921	.8257
106	.8781	.8705	.8026
107	.9170	.8719	.8295
108	.8804	.8658	.8041
109	.8763	.8523	.7930
110	.9041	.8556	.8246
Average	.8839	.8528	.8049

4   Experiences from Gathering Usable EEG Data

Given the challenges of gathering usable EEG data for valid experimental use, we now provide on overview of experiences gained from gathering the NTCIR dataset, as well as other related activities over the recent years.

Table 2. Balanced accuracy scores for each participating team's best performing method broken down by experimental task.

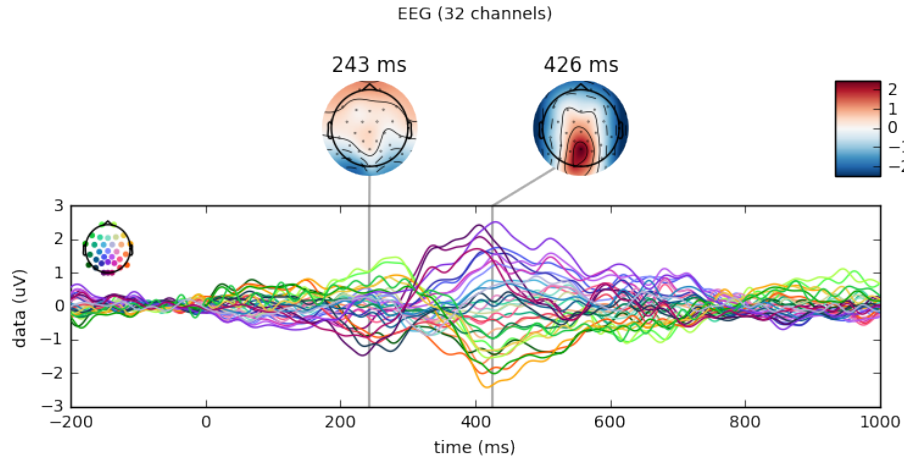| Task ID	Team-1	Team-2	Baseline
WIND1	.8905	.8609	.8237
WIND2	.8846	.8356	.7746
INSTR	.8114	.7895	.7616
BIRD	.8616	.8191	.7805
UAV1	.9216	.9053	.8381
UAV2	.9335	.9065	.8512
Average	.8839	.8528	.8049

Fig. 2. Butterfly plot (ERP averages) of target epochs across all blocks minus average standard epochs across all blocks. Plots are generated using CAR (common average reference). Characteristic P3b activity can be seen at posterior scalp sites approximately between 300ms and 600ms following target detection (peaking at 426ms). The colors on time-series plots indicate electrode location on scalp (upper left).

4.1   Data Gathering

While EEG data clearly has a lot of potential to assist the multimedia modeling, analysis and retrieval communities, recording clean EEG data is a laborious activity requiring a coordinator to strike a balance between multiple factors such as experiment length, a participant's motivation to maintain attention during the task and task complexity. There is little point in recording data if the participant is no longer engaged in a task that requires their active engagement for the data to be meaningful (as is the case with the endogenous P300 [1]).

Furthermore, if aiming for high quality signals, it requires the careful connection of an EEG cap ensuring the impedance on each electrode is kept low. This typically involves using an instrument to mechanically abrade the top layer of (dead) skin and using a conductive paste or liquid to enable conduction of electrical signals generated by the brain to the electrodes [13]. This is a messy business requiring participants after an experiment to wash their hair (necessitating some basic washing facilities on site). Moreover, care must be taken not to use too much conductive gel as electrodes can bridge together when this gel creeps along the scalp. This becomes particularly cumbersome when using more than 32 electrodes in an experiment as the electrodes will be closer together. For those new to recording EEG, this can often be daunting as there's a valid fear that they may injure a participant, which leads to over-cautious scalp preparation behaviour that in fact counter-productively results in more discomfort than necessary [2]. While many dry electrode EEG systems (not requiring a conductive gel) are available, their SNR (Signal-to-Noise Ratio) issues can often impede applications that rely on a high SNR. Due to the relatively low SNR of the P300 signal elicited in the NAILS task, a wet electrode EEG system is usually necessary.

Other practical experience learned over time from conducting similar EEG studies, which will assist researchers to gather their own EEG datasets include:

- When a participant arrives having rushed from somewhere else, they will often be sweating. Sometimes they will just be anxious or excited about the experiment causing them to sweat. It's important for the first few minutes to allow a participant to get their bearings, and this is often a good time to run through formalities such as ethics or informed consent. Sweat is deleterious to EEG signal quality and the issues it presents can often be avoided with some planning.
- Before connection to an EEG system ensure that the participant will not get disturbed or does not need to use a washroom.
- Allow a participant a practise session to get comfortable with the general requirements of the task.
- Do some pilot tests before committing to executing a large study (and have a pipeline to analyse and inspect the data from the start). It's easy to overlook what seems like a minor detail that can have deleterious effects that are only realised during data analysis when it's too late to fix a timing synchronisation or trigger problem [27].

- Check impedance of the electrodes during the experiment.
- Pay careful attention to the instructions you use to ask a participant to complete a task as they can have undesired consequences i.e. asking a participant to minimise eye blinks can often result in the subject in a RSVP experiment withholding eye blinks until they see a target image in a stream resulting in confounding eye blinks consistently occurring directly after a target image.
- Instruct the participant to try and avoid physical movements during the experimental blocks (and similarly make it apparent if the participant needs to adjust themselves physically that between experimental blocks is a good time to do so). Avoid swivel and recliner chairs as they can encourage movement during the experiment.
- Choose a quiet location to do the experiment where the participant will not be interrupted with distractions e.g. overhearing a hallway conversation distracting them from the task.
- Use small experimental blocks; in long experimental blocks it's unlikely a participant will be able to consistently sustain their attention throughout.
- EEG experiments are often long and boring, and it is often vitally important that the participant remains engaged in the task. Conversing with the participant during electrode setup and between experimental blocks will make them feel less objectified as a data point and more likely to want to do the task correctly.
- Ensure the participant is comfortable. The distraction of hunger/thirst or an uncomfortable chair (requiring regular readjustment) is not only likely to affect a participant's performance but will also likely add to giving your experiment a bad reputation when recruiting future participants.

4.2 Experiences from running NTCIR-13 NAILS

Only two of the nine teams successfully submitted an overview paper along with valid predictions to the NAILS task. Teams were contacted prior to the due date of submission of results (i.e., predictions on the test set), and the most common reason for not continuing participation was due to not achieving any significant improvements over the baseline approach. All participants in the task were provided with an implementation for the baseline approach so as to ease their participation. A table of results related to the competition are presented in Table 1. These results have been subsequently interpreted in an overview paper [7]. Elements of the NAILS task that happened at NTCIR-13, particularly the use of EEG in an informational retrieval context, will be married with a previous core task [4] as part of NTCIR-15.

Invariance of ranking of participant accuracies Although not explicitly stated in the overview paper [7], each team's winning approach (and the baseline) showed high correlation in their respective balanced accuracy scores when analysing on a per-participant basis i.e. balanced accuracy scores for individual experimental participants data tended to be similarly ranked regardless of approach. Using a Spearman's rho correlation test we find the balanced accuracy

scores for Team 1's and Team 2's best approaches significantly correlate (r_s=.86, p=.0015, N=10). This is similarly the case when performing the same correlation test against the baseline approaches for both Team 1 (r_s=.89, p=.0005, N=10) and Team 2 (r_s=.82, p=.0038, N=10). This is an important observation as it indicates some participant's EEG data was difficult to classify regardless of the explored approaches taken. When conducting a future study like this, we may actively seek to screen participants in order to collect a more directed dataset that focuses on participants who have difficult-to-classify EEG data. Similarly, we intend to explore whether it's possible to identify such experimental participants via a proxy measure like reaction time [21].

Invariance of ranking of task accuracies Performing a similar Spearman's rho correlation analysis on the balanced accuracy scores of the tasks (where scores are averaged across participants for a task) we find the scores of Team 1's and Team 2's best approaches significantly correlate (r_s=1, p=.0000, N=6). This is similarly the case when performing a similar correlation test against the baseline approaches for both Team 1 (r_s=.94, p=.0048, N=6) and Team 2 (r_s=.94, p=.0048, N=6). These results indicate that balanced accuracy scores for experimental task's data tended to be similarly ranked regardless of approach.

ARL17 [25] and QUT [9] both made successful submissions whose respective balanced accuracies on the test set are shown in Table 1 and Table 2. Both team's best results achieved balanced accuracy scores on the test set greater than any of the naive baseline approaches. This indicates that both participating team's approaches used a suitably developed strategy i.e. they outperformed a classical off-the-shelf machine-learning strategy like a SVM. In Table 1 and Table 2 we present a breakdown across participants and tasks (respectively) of the balanced accuracies achieved by each team's best performing method. Notably, independent subject nor task (or a combination thereof) models were not used as runs as it was found that these models performed sub-optimally to those trained on the data both per task and per participant. This is particularly important, as it shows there may be a need for much larger datasets to accomplish such feats.

5   Conclusions

There is a growing interest in utilising novel signal sources such as EEG (Electroencephalography) in multimedia research. While these signals can provide a useful source of evidence in multimodal media analytics, significant domain expertise is required to gather EEG datasets and make use of the resulting data. This paper presents a summary of distilled experiences and knowledge gained during the preparation (and utilisiation) of an EEG dataset that supported a collaborative neural-image labelling benchmarking task. This paper also highlights the nature of the novel EEG dataset and provides details on how it was made and how it can be accessed.

References

1. J. Y. Bennington and J. Polich. Comparison of p300 from passive and active tasks for auditory and visual stimuli. *International Journal of Psychophysiology*, 34(2):171–177, 1999.
2. J. S. Furyk, C. J. O'Kane, P. J. Aitken, C. J. Banks, and D. A. Kault. Fast versus slow bandaid removal: a randomised trial. *Medical Journal of Australia*, 191(11-12):682–683, 2009.
3. A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, et al. Meg and eeg data analysis with mne-python. *Frontiers in neuroscience*, 7:267, 2013.
4. C. Gurrin, H. Joho, F. Hopfgartner, L. Zhou, and R. Albatal. Ntcir lifelog: The first test collection for lifelog research. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 705–708. ACM, 2016.
5. G. Healy and A. F. Smeaton. Eye fixation related potentials in a target search task. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4203–4206, Aug 2011.
6. G. Healy, Z. Wang, C. Currin, T. E. Ward, and A. F. Smeaton. An eeg image-search dataset: A first-of-its-kind in IR/IIR. NAILS: neurally augmented image labelling strategies. In *In Proceedings of CHIR Workshop on Challenges in Bringing Neuroscience to Research in Human-Information Interaction*, 11 Mar 2017, Oslo, Norway.
7. G. Healy, T. E. Ward, C. Gurrin, and A. F. Smeaton. Overview of ntcir-13 nails task. 2017.
8. G. F. Healy, C. Gurrin, and A. F. Smeaton. Informed perspectives on human annotation using neural signals. In *International Conference on Multimedia Modeling*, pages 315–327. Springer, 2016.
9. H. Hutson, S. Geva, and P. Cimiano. Ensemble Methods for the NTCIR-13 NAILS Task. *In Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies, NTCIR-13, Tokyo, Japan, 5-8 December 2017.*, 2017.
10. J. W. Kam, S. Griffin, A. Shen, S. Patel, H. Hinrichs, H.-J. Heinze, L. Y. Deouell, and R. T. Knight. Systematic comparison between a wireless eeg system with dry electrodes and a wired eeg system with wet electrodes. *NeuroImage*, 184:119–129, 2019.
11. S. Koelstra, C. Mhl, M. Soleymani, J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. Deap: A database for emotion analysis using physiological signals. 3(1):18–31, 2012. eemcs-eprint-21368.
12. V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance. Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces. *Journal of neural engineering*, 15(5):056013, 2018.
13. S. J. Luck. *An introduction to the event-related potential technique.* MIT press, 2014.

14. A. R. Marathe, A. J. Ries, V. J. Lawhern, B. J. Lance, J. Touryan, K. McDowell, and H. Cecotti. The effect of target and non-target similarity on neural classification performance: a boost from confidence. *Frontiers in Neuroscience*, 9:270, 2015.

15. K. E. Mathewson, T. J. Harrison, and S. A. Kizuk. High and dry? comparing active dry eeg electrodes to active and passive wet electrodes. *Psychophysiology*, 54(1):74–82, 2017.

16. T. R. Mullen, C. A. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs. Real-time neuroimaging and cognitive monitoring using wearable dry eeg. *IEEE Transactions on Biomedical Engineering*, 62(11):2553–2567, 2015.

17. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830, 2011.

18. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

19. E. A. Pohlmeyer, J. Wang, D. C. Jangraw, B. Lou, S.-F. Chang, and P. Sajda. Closing the loop in cortically-coupled computer vision: a brain-computer interface for searching image databases. *Journal of neural engineering*, 8 3:036025, 2011.

20. J. Polich. Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol*, 118(10):2128–2148, Oct 2007.

21. A. Ramchurn, J. W. de Fockert, L. Mason, S. Darling, and D. Bunce. Intraindividual reaction time variability affects p300 amplitude rather than latency. *Frontiers in human neuroscience*, 8:557, 2014.

22. S. Razakarivony and F. Jurie. Vehicle detection in aerial imagery : A small target detection benchmark. *Journal of Visual Communication and Image Representation*, 34:187 – 203, 2016.

23. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.

24. A. F. Smeaton, O. Corrigan, P. Dockree, C. Gurrin, G. Healy, F. Hu, K. McGuinness, E. Mohedano, and T. E. Ward. Dublin's participation in the predicting media memorability task at mediaeval 2018. 2018.

25. A. J. Solon, S. M. Gordon, B. J. Lance, and V. J. Lawhern. Deep Learning Approaches for P300 Classification in Image Triage: Applications to the NAILS Task. *In Proceedings of the 13th NTCIR Conference on Evaluation of Information Access Technologies, NTCIR-13, Tokyo, Japan, 5-8 December 2017.*, 2017.

26. S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381:520 EP –, Jun 1996.

27. Z. Wang, G. Healy, A. F. Smeaton, and T. E. Ward. An investigation of triggering approaches for the rapid serial visual presentation paradigm in brain computer interfacing. In *2016 27th Irish Signals and Systems Conference (ISSC)*, pages 1–6. IEEE, 2016.

28. Z. Wang, G. Healy, A. F. Smeaton, and T. E. Ward. Use of neural signals to evaluate the quality of generative adversarial network performance in facial image generation. *Cognitive Computation*, Aug 2019.

29. B. Zhou, A. Khosla, À. Lapedriza, A. Torralba, and A. Oliva. Places: An image database for deep scene understanding. *CoRR*, abs/1610.02055, 2016.