Computer Methods and Programs in Biomedicine 221 (2022) 106929

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Machine learning algorithms on eye tracking trajectories to classify patients with spatial neglect

Benedetta Franceschiello[a,d,i,j,k,1], Tommaso Di Noto[i,1], Alexia Bourgeois[c], Micah M. Murray[a,b,d,e,j], Astrid Minier[a,b], Pierre Pouget[c], Jonas Richiardi[i,j], Paolo Bartolomeo[f], Fabio Anselmi[g,h,*]

[a] The LINE (Laboratory for Investigative Neurophysiology), Department of Diagnostic and Interventional Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland.
[b] Department of Ophthalmology, Fondation Asile des Aveugles and University of Lausanne, Lausanne, Switzerland
[c] Laboratory of Cognitive Neurorehabilitation, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[d] CIBM Center for Biomedical Imaging, Lausanne, Switzerland
[e] Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA
[f] Sorbonne Universite, Inserm, CNRS, Institut du Cerveau - Paris Brain Institute, ICM, Hopital de la Pitie-Salpetriere, Paris, France
[g] Center for Neuroscience and Artificial Intelligence, Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA
[h] Center for Brains, Minds, and Machines, McGovern Institute for Brain Research at MIT, Cambridge, MA, USA
[i] Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland
[j] The Sense Innovation and Research Center, Lausanne and Sion, Switzerland
[k] School of Engineering, Institute of Systems Engineering, HES-SO Valais-Wallis, Route de L'industrie 23, Sion, Switzerland

ARTICLE INFO

Article history:
Received 3 December 2021
Revised 19 May 2022
Accepted 31 May 2022

MSC:
0000
1111

PACS:
0000
1111

Keywords:
Neglect
Bio-markers
Eye-tracking
Machine learning
Deep networks
Structural lesion
Diffusion tensor imaging

ABSTRACT

Background and Objective: Eye-movement trajectories are rich behavioral data, providing a window on how the brain processes information. We address the challenge of characterizing signs of visuo-spatial neglect from saccadic eye trajectories recorded in brain-damaged patients with spatial neglect as well as in healthy controls during a visual search task. Methods: We establish a standardized pre-processing pipeline adaptable to other task-based eye-tracker measurements. We use traditional machine learning algorithms together with deep convolutional networks (both 1D and 2D) to automatically analyze eye trajectories. *Results:* Our top-performing machine learning models classified neglect patients vs. healthy individuals with an Area Under the ROC curve (AUC) ranging from 0.83 to 0.86. Moreover, the 1D convolutional neural network scores correlated with the degree of severity of neglect behavior as estimated with standardized paper-and-pencil tests and with the integrity of white matter tracts measured from Diffusion Tensor Imaging (DTI). Interestingly, the latter showed a clear correlation with the third branch of the superior longitudinal fasciculus (SLF), especially damaged in neglect. *Conclusions:* The study introduces new methods for both the pre-processing and the classification of eye-movement trajectories in patients with neglect syndrome. The proposed methods can likely be applied to other types of neurological diseases opening the possibility of new computer-aided, precise, sensitive and non-invasive diagnostic tools.

1. Introduction

Eye-movements are non-invasive and readily accessible behavioral readouts, providing a window onto how the brain processes information. The behavioral performance of the eyes, in particular via saccadic eye-movements, has been the focus of decades of research linking functional oculomotor behavior to dysfunction [1]. For instance, saccadic eye-movements can be a precursor of brain pathology and may also constitute an important biomarker for early diagnosis of brain impairments [2]. They may also be particularly affected after a focal brain lesion, such as in patients suffering from neglect [3–5]. Left visuo-spatial neglect (hereafter simply 'neglect') is a frequent, but still poorly understood neurological syndrome that is characterized by a lack of awareness of contrale-

* Corresponding author.
E-mail address: fabio.anselmi@bcm.edu (F. Anselmi).
[1] Equal contribution.

https://doi.org/10.1016/j.cmpb.2022.106929
0169-2607/

sional stimuli following right hemispheric damage [6]. The diagnosis of neglect is important since this syndrome is associated with poor functional outcomes [7]. A high degree of overlap between attentional orienting deficits in neglect patients and their oculomotor performance has been demonstrated. Neglect patients exhibit saccadic impairments, including direction-specific deficits of saccadic production [8,9], saccadic amplitude or difficulty retaining locations across saccades [10]. Previous studies have shown that saccadic eye-movements are a sensitive measurement to characterize neglect [11]. This is an important observation since paper-and-pencil tests administered to evaluate neglect have limited sensitivity [12]. Furthermore, the neural mechanisms underlying neglect remain debated. Neglect has been linked with structural damage of key parietal regions such as the temporo-parietal junction (TPJ) or the inferior parietal lobule [13]. Other studies indicated that damage of long-range white matter tracts, connecting frontal and parietal areas, may represent crucial antecedents of neglect [14,15]. Neglect has been reported following damage to the second and third branches of the superior longitudinal fasciculus (SLF) or to the inferior fronto-occipital fasciculus (IFOF), disrupting connectivity within attentional network of the brain [16,17].

Machine learning (ML) algorithms have been vastly employed in the analysis of medical data [18,19] due to increased need of automatising and accelerating data-analysis. In particular, convolutional neural networks (CNNs) have proven to be extremely successful machine learning architectures, for instance in supervised image classification [20]. CNNs are known to better scale with larger input datasets, when compared to traditional machine learning algorithms (e.g. kernel methods). Indeed, thanks to the enormous amount of data generated within the healthcare sector, CNNs have been effectively employed in a variety of image-driven medical diagnosis domains (e.g. radiology [21,22], ophthalmology [23]). Current research on machine learning techniques applied to eye-tracking data have hitherto focused on different fields, such as classification of eye movements (fixation, saccades, etc.) [24], or computer-assisted diagnosis tools [25]. Machine learning algorithms provide early diagnosis methods for classification and detection of neuro-developmental disorders [26], methods aiming at detecting the presence of strabismus [27], the detection of Alzheimer's disease [28,29], or of Mild Cognitive Impairment [30] based on eye-movement behavior identification. For instance, the authors in [28] use deep neural networks to identify patients with Alzheimer's disease (AD). In particular their study highlights how in principle CNNs can be used for early diagnosis of AD. Other examples of the clinical application of CNNs include for neurodevelopmental disorders [31] and strabismus [27], wherein CNNs were employed to capture subtle geometric features of eye trajectories regarding a patient's status, to which common diagnostics could be blinded. Similarly, a support vector machine algorithm successfully classified patients with memory impairments [30], or readers with dyslexia [26].

The contributions of the present work are threefold. First, the pre-processing of eye-tracker trajectories oftentimes is not detailed in its steps, despite data needing to undergo cleaning and reorganisation prior to analysis. Eye-tracking data typically contain errors and noise that must be accounted for [32]. As a first contribution, we thus provide an overview (and the corresponding code) of the pre-processing required for the task at hand, together with a standardized version of it to outlay a pipeline adaptable to other task-based eye-tracker measurements. Second, this paper demonstrates how modern machine learning algorithms, namely Support Vector Machine, Random Forests, AdaBoost and CNNs (1D and 2D), allow learning of representations of features of data that are particularly effective in classifying pathological versus non-pathological conditions from patterns of eye-movements. Our methods are situated within the growing field of automatic diagnostic tools, a branch of non-invasive techniques at the interface between neuroscience and computer science, which can transform a simple task into an automatic diagnostic procedure. Finally, we explored the anatomical correlates of eye-tracking trajectories at a network level, by combining diffusion tensor imaging (DTI) with the 1D-CNN output as predictor. To the best of our knowledge, this is the first time that machine learning methods are used to determine and quantify the presence of neglect through eye-movement analysis during a visual search task and that a link between the algorithm's outputs and anatomical markers is established.

2. Material and methods

2.1. Behavioral and neuroimaging data collection

Participants We analyzed eye movement data in a sample of 15 right-brain damaged patients with left visuo-spatial neglect and 9 healthy controls, recruited from a previous study [33]. Seven out of the 15 patients had varying degrees of left visual field defect, as assessed by confrontation testing or perimetry. Patients were considered as having significant neglect if they manifested some behavioral signs of visual neglect such as unawareness of persons or objects placed contralesionally, as well as objective signs of neglect assessed with a paper-and-pencil neglect battery. Patients diagnosed by neurologists as presenting such clinical signs of neglect could take part in the experimental session. Demographic and clinical characteristics of the patients are presented in Table 1 (1, A). Healthy individuals were age-matched with the patients (mean age 58 years, range 45–69, $t < 1$) and had no neurological or psychiatric history.

Apparatus, stimuli and procedure Participants were asked to perform a visual search task (Fig. 1) [33]. Each trial started with the presentation of a cue that lasted for 3000 ms. The cue corresponded to a white central circle surrounded by eight peripheral grey circle. The central circle was imaginarily subdivided in four quadrants; at the beginning of each trial, one of these quadrants could be filled in white, serving as an attentional cue, which indicated the most likely location of target appearance. The cue correctly indicated the target location on 73% of the trials (valid location). The target appeared in one of the three uncued quadrants (invalid location) on 18% of the trials. The target was not present on the remaining 9% of the trials (catch trials), which were included in the design to avoid guesses and anticipation. After the cue period of 3000 ms, patients were explicitly asked to maintain their gaze on the central cue and to freely move their eyes as soon as the cue disappeared. The target was presented until a manual response was made, or for 6000 ms in case of no response. The target was created by eliminating either the upper or the lower part (0.4° of visual angle) of one of the eight peripheral circles (see Fig. 1b). The remaining seven peripheral circles presented together with the target operate as distractors. The color of each target and distractors (blue, orange, red, and green) changed randomly in each trial, thus requiring an attention-demanding serial search. Participants were asked to move a joystick up when the upper part of the circle was missing, or down when the lower part was missing, as fast and as accurately as possible with their right hand. The reaction time (i.e. timestamp of the subject moving the joystick) is a psychophysical measure of the decision based on conscious perception. Therefore, in this task, errors due to involuntarily landing on the target location and resulting from the patient performing saccadic exploration are avoided by the experimental design.Eye-movements were recorded at a sampling rate of 300Hz with a Tobii TX300 eye-tracker. This experiment was composed of a total of 176 trials. *Neuroimaging data collection* Diffusion Tensor Imaging (DTI) tractography was used to study long-range of subcortical white matter pathways. For the complete pre-processing

Table 1
Classification results across the 10 random runs using the x spatial coordinate (Row 1,3,5,7), the y spatial coordinate (Row 2,4,6,8) and the 2D trajectories represented as images in the xy plane (Row 9) and concatenation of x and y coordinates into two-dimensional tensor (Row 10). Values are presented as mean ± standard deviation. SVM = Support Vector Machine; RF = Random Forest; AB = AdaBoost; Coord = spatial coordinate; Acc = accuracy; Sens = sensitivity; Spec = specificity; PPV = positive predictive value; NPV = negative predictive value; F1 = F1-score; AUC = area under the ROC curve; AUPR = area under the PR curve.

Model	Coord	Acc (%)	Sens (%)	Spec (%)	PPV (%)	NPV (%)	F1 (%)	AUC	AUPR
1D-CNN	x	86 5	80 11	89 3	82 6	88 6	81 8	.85 0.06	.85 0.06
1D-CNN	y	85 3	79 6	88 4	80 5	88 3	79 4	.83 0.03	.83 0.03
SVM	x	88 4	72 6	97 3	93 7	85 3	81 6	.84 0.04	.88 0.05
SVM	y	77 2	70 5	81 4	70 4	82 2	70 3	.76 0.02	.75 0.02
RF	x	88 3	74 5	96 4	92 8	86 2	82 5	.85 0.04	.88 0.05
RF	y	87 2	78 0	92 3	86 4	87 0	81 2	.85 0.01	.86 0.02
AB	x	89 3	76 7	97 3	95 6	87 3	84 5	.86 0.04	.90 0.04
AB	y	87 4	81 5	90 5	83 7	89 3	82 5	.86 0.04	.86 0.05
CNN 2D-image	xy	78 5	72 7	82 7	71 8	83 4	71 6	.77 0.05	.77 0.05
CNN 2D-vector	xy	85 5	83 15	87 5	79 5	91 8	80 8	.85 0.07	.84 0.05

pipeline, see [33]. The mean fractional anisotropy values of the three branches of the superior longitudinal fasciculus (SLF), the cingulum, and the inferior fronto-occipital fasciculus (IFOF) were extracted in the right hemisphere. These tracts were chosen on the basis of their implication in attention networks and in visual neglect [14]. The analysis was conducted on n=13 patients, because the MRI scans were not available for two patients.

2.2. Data processing and analysis

Pre-processing Targets presented within the left visual field were considered for the analysis, as neglect symptoms concern first and foremost attentional orienting towards the left, contralesional hemispace. Most studies report that neglect patients show distinct eye movement patterns, with the majority of studies indicating that more eye movements were made towards the ipsilesional right side of space [35]. However, in our previous study [33], specific patterns of eye-movements were observed when left-sided targets were presented. We thus aimed at quantifying these left side trajectories which could represent a unique signature in neglect. Recorded eye-tracking trajectories underwent a pre-processing procedure to standardize the dataset before analyzing it. The code was developed with Matlab R2019b, see paragraph 2.2. First, the duration of the trajectories was standardized across trials, as the following machine learning models must have same-length vectors as input. All trials - those where the participant replied with the joystick and those were the participant could not reach the target - were uniformed in order to be 9000 ms long. Time series corresponding to trials ending before 9*s* were filled with NaN (not a number, missed recording), for a pure pre-processing reason. An actual number was attributed in a second phase of the pre-processing. The analysis focused on those trials both valid, invalid and not missed by the subject, according to the definition given in Section 2.1. As the eye-tracker continuously acquires across time and we were interested in the visual search part of the protocol, only time-stamps recorded after the first 3000 ms and before 9000 ms were first taken into account. A loop over the total number of targets seen by each subject was iterated to analyze all trajectories recorded per each participant. For each target i the set of trajectories' coordinates $(Eye_x, Eye_y)_i$ was extracted and coordinates corresponding to outliers (such as points outside the screen or above 9000 ms) were respectively filled with NaN if acquired between 3000 and 9000 ms, otherwise removed. All missed recordings were then re-filled according to the following criteria: if a NaN was present at the beginning of the trial, the NaN value was replaced with the center of the screen coordinates (i.e. at pixels $(384, 512)$). If a NaN was in the middle of the time series, this value was interpolated (nearest neighbor) using the available neighboring sampled data, corresponding to actual recordings. Nearest neighbor was chosen to imitate the non-continuous nature of saccades in this experiment, where the participants were asked to follow the presentation of stimuli in different loci. If the target was reached at a certain timestamp, all remaining points (to reach the 9000 ms upper limit) were filled with the target coordinates. The first 3000 ms of Eye_x and Eye_y was erased, as they corresponded to the fixation part, leaving us with a shortened time series 6000 ms long, now cleaned and interpolated. Also, the latter 3000 ms were eliminated to avoid piece-wise constant trajectories, as participants in the considered trials were reaching the target beforehand (see Fig. 1). This avoided injecting a constant piece of the trajectory in the network, as the goal here was to assess eye-trajectory variability. The post-processed time-series correspond to the first 3000ms of the visual search task. Every trial for every subject and every target i, $(Eye_x, Eye_y)_i$, was z-scored by subtracting its mean and dividing by its standard deviation. All pre-processing steps are schematized in Fig. 1. Only targets presented to the left visual field of the participant were considered. To extend this pre-processing pipeline to other oculomotor experiments one should consider applying the following steps: 1) standardize the vectors length for the various ML models employed, 2) flag the outliers, 3) treat the missing recording and outliers via interpolation - preferably nearest neighbour to preserve the eye behavior, 4) discard constant tales at the end of the vector, and 5) normalize each trajectory.

Classifying healthy vs. neglect patients from eye-tracking trajectories

In this section, we discuss the methodology we applied to estimate healthy versus neglect patients' status from a subset of eye-tracking trajectories, i.e. those corresponding to targets within the left hemispace. Left hemispace targets are more challenging for patients affected by this syndrome. For this purpose, we formulated the estimation problem as a classification task, i.e. learning the mapping between an appropriate representation of the eyes' trajectories and the patient's label. In particular, we separately analyzed the x coordinates and y coordinates of the trajectories (x-projection and y-projection), of targets presented to the left visual field of the participant. The resulting vectors have length $d = 1001$ and are used as input samples to our classifiers. We fed the 1D trajectories to several machine-learning classifiers, namely Support Vector Machine (SVM), Random Forest (RF), AdaBoost (AB) and a 1D-CNN. In addition, since the eye movements are intrinsically 2-dimensional, we also investigated the performances of our CNN when trained both on 2D images representing the trajectories in the xy plane (henceforth called CNN 2D-image) and on a two-dimensional tensor consisting of a concatenation of the x and y coordinates (henceforth called CNN 2D-vector). In the former sce-

A

	Sex, age	Onset of Illness (days)	Aetiology	Visual field	Bells cancellation (left/right hits)	Letter cancellation (Left, right hits)	Line bisection (mm of rightward deviation for 200mm)	Landscape drawing	Reading (Left/right hits)
1	M, 60+	115	I	Intact	12/15*	25/30*	5,5	0	61/55
2	F, 60+	210	I	Intact	0/7*	NA	5,5	3*	0/12*
3	M, 50+	431	I	Intact	14/14	23/28*	16*	0	61/55
4	M, 60+	395	I	Intact	12/15*	24/28*	0,1	3*	59/55
5	F, 50+	415	H	Intact	14/13	30/29	12.5*	5*	61/55
6	M, 60+	239	H	Intact	10/13*	NA	9.5*	0	58/55*
7	M, 40+	745	H	Intact	09/15*	30/30	9.5*	0	61/55
8	F,60+	190	I	Hemianopia	12/15*	20/30*	6	0	46/55*
9	M,60+	700	I	Hemianopia	15/14	25/25	4	0	55/55*
10	F,40+	132	I/H	Hemianopia	0/12*	0/21*	50*	3*	25/54*
11	F,60+	89	I	Hemianopia	1/11*	6/26*	53.5*	4*	2/35*
12	M,50+	106	I	Hemianopia	12/15*	28/29	0,2	0	59/55*
13	F,50+	253	I	Hemianopia	0/6*	NA	38.5*	4*	0/37*
14	M,60+	264	I	Hemianopia	04/10*	NA	21*	1*	57/55*
15	M,70+	313	I	Hemianopia	09/15*	NA	10*	1*	56/55*

Cut-off scores: Bells cancellation, difference between left and right omissions >2 (Azouvi et al., 2006); Letter cancellation, one target in each field can rest undetected (Mesulam, 1985), Line bisection, -7.3 mm for leftward deviation, 6.5 mm for rightward deviation (Azouvi et al., 2006); Landscape drawing scores >0 (Azouvi et al., 2006); Reading, difference between left and right omissions >0 (Azouvi et al., 2006).

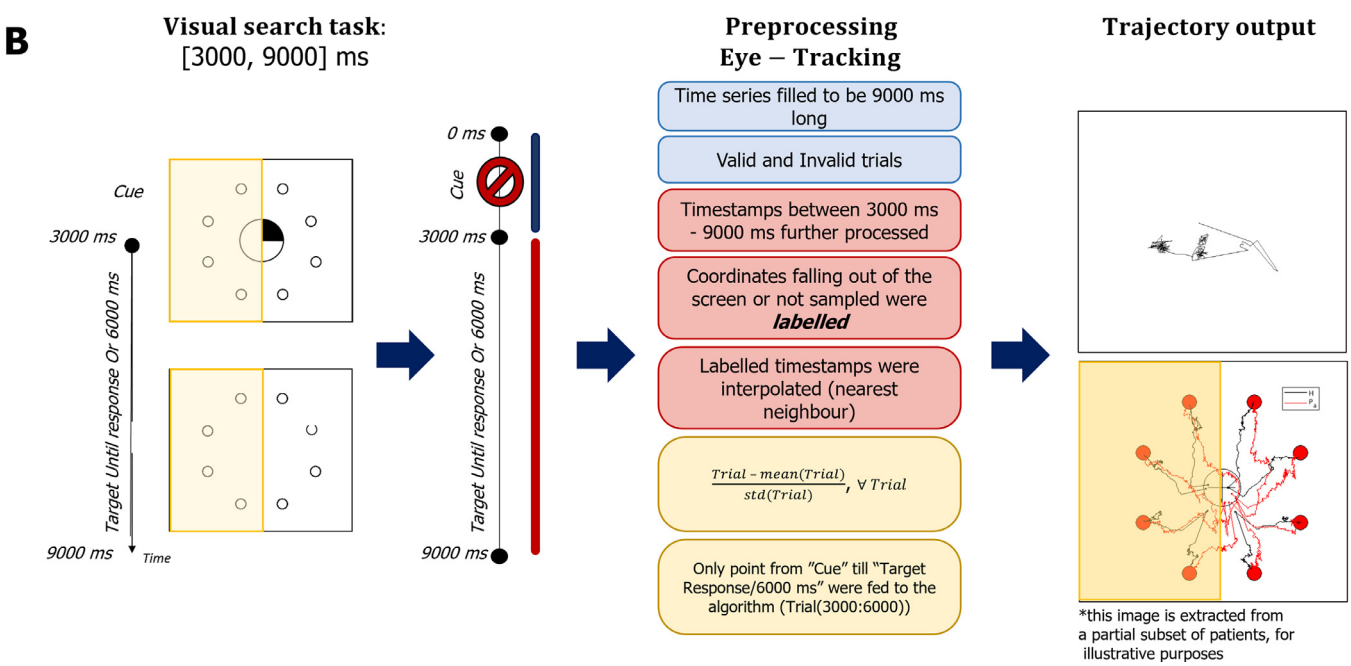

Fig. 1. (Top). A. Demographic and clinical characteristics of neglect patients, with their performance on visuo-spatial tests. Asterisks denote pathological performance compared to normative data. For line bisection, positive values indicate rightward deviations, negative values indicate leftward deviations. Scores for the landscape drawing [34] indicate the number of omitted left-sided details. I, ischemic; H, hemorrhagic; NA, not available. B. The pre-processing steps required to analyse the eye-tracker data and the corresponding parts of the visual search behavioral paradigm task they refer to. (Bottom-right) An example of pre-processed trial, as well as the mean trajectories over all trials of a subset of subjects for each target. In the bottom row the left yellow box refers to the left-sided targets used for the analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nario, we cropped the images around the non-zero xy coordinates to reduce sparsity (by using the full images the network did not converge) and then we zero-padded them to have a uniform image shape across all samples.

For the traditional ML classifiers (i.e. SVM, RF and AB), we used the vanilla implementations from scikit-learn. Instead, for the 1D (and 2D) CNNs, we built a custom model with building blocks inspired by the VGG-16 network [36]. Specifically, our network is composed of a sequence of 3 blocks of convolutional layers, each one separated by a pooling layer that halves the output vector dimension. A kernel size of 3 and zero padding were used in all convolutional layers. To obtain the desired classification output, the convolutional blocks are followed by three dense, fully connected (FC) layers. We applied the ReLU activation function for all layers, except for the last layer which is followed by a sigmoid function. We used Xavier initialization [37] for all layers. Biases were initialized to 0 and a batch size of 16 was chosen. Batch normalization and dropout (rate = 0.3) were used to avoid overfitting in the

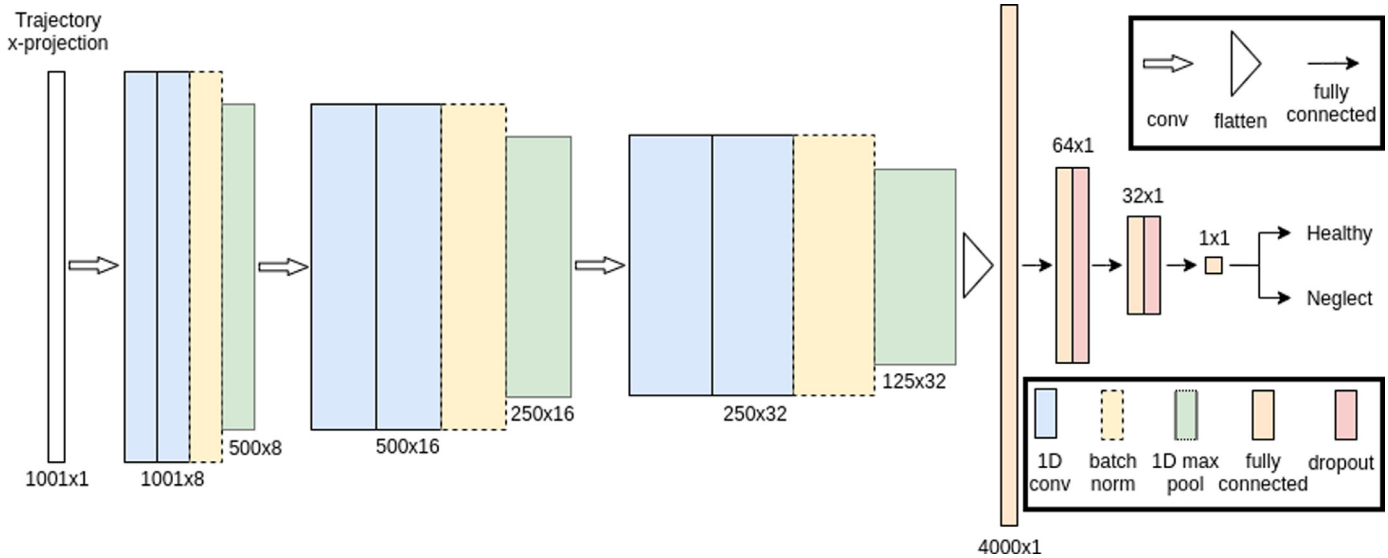

Fig. 2. The 1D-CNN architecture used in our experiments. Three convolutional blocks with batch normalization and pooling are followed by two dense layers with dropout. Dimensions of each layer are reported in the image. The final output is the probability that each input trajectory has to either belong to a healthy control or to a patient with spatial neglect. To assign the final class (healthy vs. neglect) to one subject, we performed a majority voting across all trajectories of that subject.

convolutional blocks and in the FC layers, respectively. To fit the model, the Adam optimization algorithm [38] was applied, with a learning rate of 0.0001. We trained the model for 25 epochs and we adopted the binary cross-entropy loss function. The total number of trainable parameters in our network is 264,377. Training and inference were implemented using Tensorflow (version 2.5). The detailed structure of our 1D-CNN is illustrated in Fig. 2. For the 2D experiments (CNN 2D-image and CNN 2D-vector), we simply converted the 1D-CNN to 2D without any modifications to the order of the layers (further details about the network implementation can be found in the github repository - see paragraph *Code and Dataset release*).

To compute unbiased test results, we applied a 5-fold cross-validation (CV) at the individual participant level. This ensured that multiple trajectories of the same participant did not appear both in the test and training set. Furthermore, to account for the variability introduced by the random choice of patients at each CV split, the whole CV was repeated 10 times (10 runs), each time performing the splitting anew, and results were averaged. The CV strategy partially remedies the lack of a large and more heterogeneous population of patients, by simulating an untouched test set and controlling for possible overfitting of the model. To statistically compare the different classifiers (e.g. 1D-CNN vs. SVM, or 1D-CNN vs. CNN 2D-image vs. CNN 2D-vector), we performed a Wilcoxon signed-rank test [39] of the AUC values across the 10 random runs, setting a significance threshold level $\alpha = 0.05$.

Finally, we computed the confidence score for each participant. This corresponds to the ratio between the correctly classified trajectories and the total number of trajectories. It can be interpreted as a measure of confidence c with respect to the final prediction. For instance, $c = .99$ implies an extremely confident correct classification; $c = .65$ implies a moderately confident correct classification; $c < .5$ implies a wrong classification. To obtain one single confidence score per participant, we averaged the confidence scores of each participant across the 10 random runs. The developed code can be adapted to the analysis of visual search tasks and possibly integrated in existing algorithms, such as those identifying fixation and regions of interest of eye-tracker trajectories [40] as well as those assessing data quality [41].

Code and Dataset release In compliance with open-source practices, we released a task-dependent version of the code used for the analyses (pre-processing and machine-learning classification) here:

https://github.com/bfranceschiello/EyeTracking_preprocessing_and_ML_analysis. The full dataset used for the analysis can be found at the following link: https://zenodo.org/record/6424677.

3. Results

Behavioral results In [33], oculomotor behaviour was studied together with manual responses on a visual search task in neglect to determine the relation between saccadic parameters and sub-clinical disorders of spatial attention. The study showed the occurrence of inappropriate rightward saccades during target selection: when left-sided targets were presented, saccades were equally likely to be performed towards the left side or towards the right hemispace. Right-sided distractors may erroneously capture patients' attention, leading to an over-exploration of the right hemispace, consistent with the so-called magnetic attraction of gaze typically observed in neglect [42]. Thus, pathological production of eye movements should be considered as a subtle manifestation of visuo-spatial disorders.

Classification results Rows 1 to 8 of Table 1 illustrate the classification test results averaged across the 10 random runs when using the x and the y projections, respectively for the various ML models employed. Row 9 reports the performance of the CNN 2D-image experiment. Row 10 reports the performance of the CNN 2D-vector experiment. To assess whether one of these methods had significantly higher performance, we compared the AUCs of the random runs with several Wilcoxon signed-rank tests. Overall, classification results using the x-coordinate consistently proved to be higher across all methods, a finding that is in agreement with the observations in [3]. For this reason, we only compared models for the x-coordinate. We found no statistically-significant difference for the AUC distributions of 1D-CNN vs. SVM ($W = 15$, $p = 0.37$). Similarly, we found no statistically-significant difference when comparing 1D-CNN vs. Random Forest ($W = 15.5$, $p = 0.40$) or the 1D-CNN vs. AdaBoost ($W = 17$, $p = 0.88$). These results indicate that all models perform equally well when using the x-projection as input.

The results we obtained are in line with both a technical and clinical perspective. On the one hand, as mentioned in Section 1,

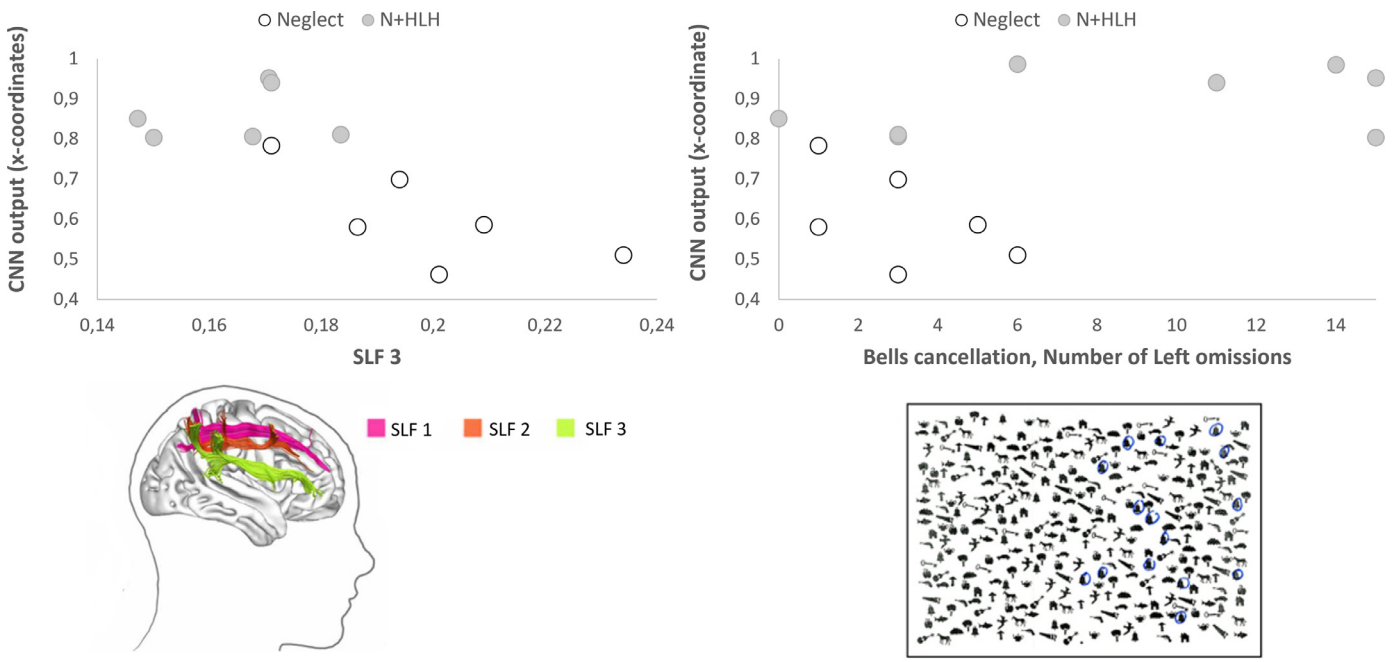

Fig. 3. The correlation between the 1D-CNN algorithm confidence scores and the Fractional Anisotropy (FA) of the SLF3 is presented (left panel) ($\rho = -0.77$), as well as the correlation between the algorithm confidence scores and the number of left omissions of the Bells cancellation test (right panel) ($\rho = 0.55$, $p = 0.033$). The images depicting the SLF and the Bells test are reprinted from Bartolomeo et al. [43].

patients' healthy vs. unhealthy status can indeed be evaluated by extracting salient features from the geometry of eye-trajectories in visual tasks, using ML approaches. On the other hand, neglect features can also be equally well identified either from x or y trajectories [3], although the performance is more efficient on x-ones. In the case of CNNs, when comparing the 1D-CNN (Row 1, Table 1) with the CNN 2D-image (Row 9, Table 1) through the Wilcoxon test, we found that AUCs of the CNN 2D-image were significantly lower than those of the 1D-CNN ($W = 2.5$, p=0.01), indicating that the 1D coordinates are more informative to distinguish healthy controls from patients with spatial neglect. Conversely, when comparing the 1D-CNN (Row 1, Table 1) with the CNN 2D-vector (Row 10, Table 1), we found no significant difference for the AUCs. This indicates that the 1D x coordinates are already discriminative enough for the task at hand.

Relationship between 1D-CNN results and Neuroimaging

Given the equal performance of the 1D ML methods and our intention to generalize the analysis to a larger dataset, we assessed the correlation between the anatomical markers of the patients and the model output only for the 1D-CNN. We performed Pearson correlations between the 1D-CNN algorithm confidence score on the x and y-spatial coordinate and the FA of long-range of white matter tracts connecting attentional networks. One patient was considered as an outlier and was discarded from the analysis. Such investigation is related to a previous study [33], where a significant correlation between an abnormal saccadic exploration behavior and impairments of the right SLF2 was found. This suggests that oculomotor behavior might be a better indicator of neglect signs than paper-and-pencil tests alone, which require more voluntary top-down orienting of attention and which may partly lead to a compensation of neglect-related deficits. Our results show a negative correlation ($\rho = -0.77$) between the 1D-CNN algorithm confidence score on the x-spatial coordinate and damage of the SLF3 ($p = 0.003$ Bonferroni-corrected, see Fig. 3, left panel). The correlations with the others tracts (SLF1, SLF2, IFOF, Cingulum), as well as between the y-coordinate and all the other tracts did not reach significance (all Bonferroni- corrected $p > 0.011$). Also, the 1D-CNN algorithm score on the x-spatial coordinate correlated with the number of left omissions in Bells cancellation [44]; a standardized paper-and-pencil test use to diagnose neglect signs ($\rho = 0.55$, $p = 0.033$) 3, right panel).

4. Discussion and conclusions

Neglect is a multi-component syndrome; dissociated performance on different tests is often observed both between and within patients. Some of these dissociations may depend on the activity of compensatory mechanisms, such as top-down orienting of attention [45], which is perhaps partly driven by the healthy left hemisphere [46]. Sensitive behavioral techniques, such as manual response times or eye movements characterization, [3–5,47,48] may thus greatly help diagnosis of neglect.

Our work demonstrates how machine learning algorithms allow learning of representations of saccadic eye-movements' features that are particularly effective in classifying neglect versus healthy conditions with AUCs in the range of 0.83 to 0.86. Overall, all employed machine learning algorithms performed equally well when using the x-projection of the trajectories. However, in the paper we focused more on the CNN algorithm, because when given a sufficiently large dataset as input, CNNs are capable of learning more discriminative features from the input data with respect to more traditional algorithms (like kernel machines). Future analyses will involve a larger input dataset to exploit richer features within the data and the CNN's scalability. Our results highlight that, a 1D-CNN trained on x projections of the eye trajectories performs better then a CNN trained on 2D images containing the trajectories. However, the same 1D-CNN performs on par with a CNN trained on two-dimensional tensors made of the combined x and y coordinates. To the best of our knowledge, this is the first time that ML methods, and in particular CNNs, are used to determine the presence of neglect through eye-movement analysis during a visual search task.

The correlation between the (1D-CNN) algorithm confidence score output and the anatomical attributes of the patients' DTI

benchmark the relevance of the technique and its specificity in detecting neglect patients. Furthermore, the 1D-CNN algorithm confidence score on the *x*-coordinates of the saccade trajectories appears to be related to the disruption of the third branch of the superior longitudinal fasciculus (SLF3). The SLF3 links parietal and frontal regions and has been shown to be specifically impaired in neglect [49,50]. We also observed a correlation between the 1D-CNN algorithm confidence score on the *x*-coordinates and the degree of severity of neglect signs ($\rho = 0.55,\ p = 0.033$). Although the algorithm confidence scores are similar for the *x* and *y*-projections of the saccades, the correlations with the anatomical and behavioral data show that predominant features seem to be contained in the *x*-projections.

Despite several studies that have highlighted pathological eye-movement behavior as a consequence of impaired shifts of attention in neglect, the measurement of eye-movement behavior as a tool to diagnose neurological pathology has only begun to be developed. In this respect our work therefore represents a contribution to understand and predict the trajectory of individual patients, opening the possibility of a new computer-aided diagnosis and follow-up tool for neglect syndrome. Further investigations should consolidate this link, allowing to differentiate and predict patterns in agreement with the anatomical markers with unprecedented precision, especially as the neural substrates of neglect are still debated, despite this syndrome representing a unique opportunity to underpin the underlying mechanisms of spatial processing and conscious awareness. One limitation of the present study comes from the lack of a larger and more heterogeneous population of patients. In the future, we plan to extend our efforts in a larger pool of subjects. In this respect, a CNN algorithm will be our preferred methods given its scalability to large datasets. Finally, this work represents a first step towards the use of machine learning techniques to other neurological conditions characterized by impaired eye movements: a growing and exciting field of investigation.

Declaration of Competing Interest

The authors certify that they have no financial and personal relationships with other people or organisations that could inappropriately influence the present work. Manuscript title: “Machine learning algorithms on eye tracking trajectories to classify patients with spatial neglect” The authors (Benedetta Franceschiello, Tommaso Di Noto, Alexia Bourgeois, Micah M. Murray, Astrid Minier, Pierre Pouget, Jonas Richiardi, Paolo Bartolomeo, Fabio Anselmi) certify that they have no financial and personal relationships with other people or organisations that could inappropriately influence the present work.

Acknowledgments

B. F. and A. M acknowledge the support of the Fondation Asile des aveugles. M. M. M. acknowledges the support of grants from the Swiss National Science Foundation (grant 320030-169206), the Fondation Asile des aveugles, and a grantor advised by Carigest SA. The Sense Innovation and Research Center is supported by the Lausanne University Hospital Center (CHUV), The University of Lausanne (UNIL), and The University of Applied Sciences of Western Switzerland (HES-SO Valais/Wallis). The work of P.B. is supported by the Agence Nationale de la Recherche through ANR-16-CE37-0005 and ANR-10-IAIHU-06, and by the Fondation pour la Recherche sur les AVC through FR-AVC-017. F. A. acknowledges the Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216 and the financial support of the AFOSR projects FA9550-17-1-0390 and BAA-AFRL-AFOSR-2016-0007 (European Office of Aerospace Research and Development), and the EU H2020-MSCA-RISE project NoMADS - DLV-777826. The authors would like to thank the anonymous Reviewers, as their comments greatly improved the quality of the manuscript. B.F. would like to thank Ms. Alicja Olszewska and Ms. Monika Puchalska for their valuable work on the repository for generalised Eye-Tracking preprocessing, developed together within the context of BrainHack Warsaw 2022. This generalized version of the pre-processing code is available here: https://github.com/bfranceschiello/ET_Analysis_BrainHack_WorkInProgress. B.F. acknowledges the organisers of BrainHack Warsaw 2022 for the organisation and resources made available during the Hackathon. Such events greatly benefit the community via open source code practises, in this case related to the scope of the article.

References

[1] T.J. Anderson, M.R. MacAskill, Eye movements in patients with neurodegenerative disorders, Nat. Rev. Neurol. 9 (2) (2013) 74.
[2] R.J. Leigh, D.S. Zee, The Neurology of Eye Movements, Contemporary Neurology, 2015.
[3] R.M. Müri, D. Cazzoli, T. Nyffeler, T. Pflugshaupt, Visual exploration pattern in hemineglect, Psychol. Res. 73 (2) (2009) 147–157.
[4] M. Behrmann, T. Ghiselli-Crippa, I. Dimatteo, Impaired initiation but not execution of contralesional saccades in hemispatial neglect, Behav. Neurol. 13 (1, 2) (2001) 39–60.
[5] E. Natale, C.A. Marzi, E. Bricolo, L. Johannsen, H.-O. Karnath, Abnormally speeded saccades to ipsilesional targets in patients with spatial neglect, Neuropsychologia 45 (2) (2007) 263–272.
[6] P. Bartolomeo, Attention Disorders after Right Brain Damage: Living in Halved Worlds, Springer, 2013.
[7] M. Jehkonen, M. Laihosalo, J.E. Kettunen, Impact of neglect on functional outcome after stroke–a review of methodological issues and recent research findings, Restor. Neurol. Neurosci. 24 (4–6) (2006) 209–215.
[8] M. Behrmann, S. Watt, S.E. Black, J. Barton, Impaired visual search in patients with unilateral neglect: an oculographic analysis, Neuropsychologia 35 (11) (1997) 1445–1458.
[9] R. Walker, J.M. Findlay, Saccadic eye movement programming in unilateral neglect, Neuropsychologia 34 (6) (1996) 493–508.
[10] M. Husain, S. Mannan, T. Hodgson, E. Wojciulik, J. Driver, C. Kennard, Impaired spatial working memory across saccades contributes to abnormal search in parietal neglect, Brain 124 (5) (2001) 941–952.
[11] B.C. Kaufmann, D. Cazzoli, T. Pflugshaupt, S. Bohlhalter, T. Vanbellingen, R.M. Müri, T. Nef, T. Nyffeler, Eyetracking during free visual exploration detects neglect more reliably than paper-pencil tests, Cortex 129 (2020) 223–235.
[12] P. Azouvi, P. Bartolomeo, J.-M. Beis, D. Perennou, P. Pradat-Diehl, M. Rousseaux, A battery of tests for the quantitative assessment of unilateral neglect, Restor. Neurol. Neurosci. 24 (4–6) (2006) 273–285.
[13] M. Corbetta, G.L. Shulman, Spatial neglect and attention networks, Annu. Rev. Neurosci. 34 (2011) 569–599.
[14] P. Bartolomeo, M. Thiebaut de Schotten, F. Doricchi, Left unilateral neglect as a disconnection syndrome, Cereb. Cortex 17 (11) (2007) 2479–2490.
[15] S. Clarke, S. Crottaz-Herbette, Modulation of visual attention by prismatic adaptation, Neuropsychologia 92 (2016) 31–41.
[16] M.T. de Schotten, M. Urbanski, H. Duffau, E. Volle, R. Lévy, B. Dubois, P. Bartolomeo, Direct evidence for a parietal-frontal pathway subserving spatial awareness in humans, Science 309 (5744) (2005) 2226–2228.
[17] M. Thiebaut de Schotten, F. Tomaiuolo, M. Aiello, S. Merola, M. Silvetti, F. Lecce, P. Bartolomeo, F. Doricchi, Damage to white matter pathways in subacute and chronic spatial neglect: a group study and 2 single-case studies with complete virtual in vivo tractography dissection, Cereb. Cortex 24 (3) (2014) 691–706.
[18] A. Dhillon, A. Singh, Machine learning in healthcare data analysis: a survey, J. Biol. Today's World 8 (6) (2019) 1–10.
[19] K.Y. Ngiam, W. Khor, Big data and machine learning algorithms for health-care delivery, Lancet Oncol. 20 (5) (2019) e262–e273.
[20] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[21] M. Cicero, A. Bilbily, E. Colak, T. Dowdell, B. Gray, K. Perampaladas, J. Barfett, Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs, Invest. Radiol. 52 (2016), doi:10.1097/RLI.0000000000000341.
[22] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D.D.F. Feng, D. Liang, Accelerating magnetic resonance imaging via deep learning, 2016, pp. 514–517, doi:10.1109/ISBI.2016.7493320.
[23] R. Poplin, A. Varadarajan, K. Blumer, Y. Liu, M. McConnell, G. Corrado, L. Peng, D. Webster, Predicting cardiovascular risk factors from retinal fundus photographs using deep learning, Nat. Biomed. Eng. 2 (2018), doi:10.1038/s41551-018-0195-0.
[24] R. Zemblys, D.C. Niehorster, O. Komogortsev, K. Holmqvist, Using machine learning to detect events in eye-tracking data, Behav. Res. Methods 50 (1) (2018) 160–181.

[25] N. Khosravan, H. Celik, B. Turkbey, E.C. Jones, B. Wood, U. Bagci, A collaborative computer aided diagnosis (c-CAD) system with eye-tracking, sparse attentional model, and deep learning, Med. Image Anal. 51 (2019) 101–115.
[26] L. Rello, M. Ballesteros, Detecting readers with dyslexia using machine learning with eye tracking measures, in: Proceedings of the 12th Web for All Conference, 2015, pp. 1–8.
[27] Z. Chen, H. Fu, W.-L. Lo, Z. Chi, Strabismus recognition using eye-tracking data and convolutional neural networks, J. Healthc Eng. 2018 (2018).
[28] J. Biondi, G. Fernandez, S. Castro, O. Agamennoni, Eye-movement behavior identification for AD diagnosis, arXiv preprint arXiv:1702.00837(2017).
[29] J. Beltrán, M.S. García-Vázquez, J. Benois-Pineau, L.M. Gutierrez-Robledo, J.-F. Dartigues, Computational techniques for eye movements analysis towards supporting early diagnosis of Alzheimers disease: a review, Comput. Math. Methods Med. 2018 (2018).
[30] D. Lagun, C. Manzanares, S.M. Zola, E.A. Buffalo, E. Agichtein, Detecting cognitive impairment by eye movement analysis using automatic classification algorithms, J. Neurosci. Methods 201 (1) (2011) 196–203.
[31] G. Pusiol, A. Esteva, S.S. Hall, M. Frank, A. Milstein, L. Fei-Fei, Vision-based classification of developmental disorders using eye-movements, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 317–325.
[32] K. Holmqvist, M. Nyström, F. Mulvey, Eye tracker data quality: what it is and how to measure it, in: Proceedings of the Symposium on Eye Tracking Research and Applications, 2012, pp. 45–52.
[33] A. Bourgeois, A.B. Chica, R. Migliaccio, D.J. Bayle, C. Duret, P. Pradat-Diehl, M. Lunven, P. Pouget, P. Bartolomeo, Inappropriate rightward saccades after right hemisphere damage: oculomotor analysis and anatomical correlates, Neuropsychologia 73 (2015) 1–11.
[34] G. Gainotti, P. D'erme, D. Monteleone, M.C. Silveri, Mechanisms of unilateral spatial neglect in relation to laterality of cerebral lesions, Brain 109 (4) (1986) 599–612.
[35] D.D. Cox, R.L. Savoy, Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex, Neuroimage 19 (2) (2003) 261–270. http://www.fmri.columbia.edu/pdfs/BrainReadingAsPublished.pdf
[36] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).
[37] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, JMLR Workshop and Conference Proceedings, 2010, pp. 249–256.
[38] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980(2014).
[39] F. Wilcoxon, Individual comparisons by ranking methods, in: Breakthroughs in statistics, Springer, 1992, pp. 196–202.
[40] R. Caldara, S. Miellet, iMap: a novel method for statistical fixation mapping of eye movement data, Behav. Res. Methods 43 (3) (2011) 864–878.
[41] D.R. van Renswoude, M.E.J. Raijmakers, A. Koornneef, S.P. Johnson, S. Hunnius, I. Visser, Gazepath: an eye-tracking analysis tool that accounts for individual differences and data quality, Behav. Res. Methods 50 (2) (2018) 834–852.
[42] G. Gainotti, P. D'Erme, P. Bartolomeo, Early orientation of attention toward the half space ipsilateral to the lesion in patients with unilateral brain damage, J. Neurol. Neurosurg. Psychiatry 54 (12) (1991) 1082–1089.
[43] P. Bartolomeo, M. Thiebaut De Schotten, A.B. Chica, Brain networks of visuospatial attention and their disruption in visual neglect, Front. Hum. Neurosci. 6 (2012) 110.
[44] L. Gauthier, F. Dehaut, Y. Joanette, et al., The bells test: a quantitative and qualitative test for visual neglect, Int. J. Clin. Neuropsychol. 11 (2) (1989) 49–54.
[45] P. Bartolomeo, E. Sieroff, C. decaix, S. Chokron, Modulating the attentional bias in unilateral neglect: the effects of the strategic set, Exp. Brain Res. 137 (2001) 432–444, doi:10.1007/s002210000642.
[46] P. Bartolomeo, From competition to cooperation: visual neglect across the hemispheres, Rev. Neurol. 177 (2021), doi:10.1016/j.neurol.2021.07.015.
[47] P. Bartolomeo, The novelty effect in recovered hemineglect, Cortex 33 (2) (1997) 323–333.
[48] P. Bartolomeo, Inhibitory processes and spatial bias after right hemisphere damage, Neuropsychol. Rehabil. 10 (5) (2000) 511–526.
[49] M. Lunven, P. Bartolomeo, Attention and spatial cognition: neural and anatomical substrates of visual neglect, Ann. Phys. Rehabil. Med. 60 (3) (2017) 124–129.
[50] M. Urbanski, M.T. De Schotten, S. Rodrigo, C. Oppenheim, E. Touzé, J.-F. Méder, K. Moreau, C. Loeper-Jeny, B. Dubois, P. Bartolomeo, DTI-MR tractography of white matter damage in stroke patients with neglect, Exp. Brain Res. 208 (4) (2011) 491–505.