J Neurol (2013) 260:275–284
DOI 10.1007/s00415-012-6631-2

ORIGINAL COMMUNICATION

High-throughput classification of clinical populations from natural viewing eye movements

Po-He Tseng · Ian G. M. Cameron · Giovanna Pari · James N. Reynolds · Douglas P. Munoz · Laurent Itti

Received: 2 May 2012 / Revised: 17 July 2012 / Accepted: 18 July 2012 / Published online: 25 August 2012

Abstract Many high-prevalence neurological disorders involve dysfunctions of oculomotor control and attention, including attention deficit hyperactivity disorder (ADHD), fetal alcohol spectrum disorder (FASD), and Parkinson's disease (PD). Previous studies have examined these deficits with clinical neurological evaluation, structured behavioral tasks, and neuroimaging. Yet, time and monetary costs prevent deploying these evaluations to large at-risk populations, which is critically important for earlier detection and better treatment. We devised a high-throughput, low-cost method where participants simply watched television while we recorded their eye movements. We combined eye-tracking data from patients and controls with a computational model of visual attention to extract 224 quantitative features. Using machine learning in a workflow inspired by microarray analysis, we identified critical features that differentiate patients from control subjects. With eye movement traces recorded from only 15 min of videos, we classified PD versus age-matched controls with 89.6 % accuracy (chance 63.2 %), and ADHD versus FASD versus control children with 77.3 % accuracy (chance 40.4 %). Our technique provides new quantitative insights into which aspects of attention and gaze control are affected by specific disorders. There is considerable promise in using this approach as a potential screening tool that is easily deployed, low-cost, and high-throughput for clinical disorders, especially in young children and elderly populations who may be less compliant to traditional evaluation tests.

Keywords ADHD · FASD · Parkinson's disease · Attention deficits · Eye tracking

Electronic supplementary material The online version of this article (doi:10.1007/s00415-012-6631-2) contains supplementary material, which is available to authorized users.

P.-H. Tseng · L. Itti (✉)
Department of Computer Science, University of Southern California, Los Angeles, CA 90089, USA
e-mail: itti@usc.edu
URL: http://iLab.usc.edu

I. G. M. Cameron · G. Pari · J. N. Reynolds · D. P. Munoz
Centre for Neuroscience Studies, Queen's University, Kingston, ON K7L 3N6, Canada

G. Pari · D. P. Munoz
Department of Medicine, Queen's University, Kingston, ON K7L 3N6, Canada

J. N. Reynolds · D. P. Munoz
Department of Biomedical and Molecular Science, Queen's University, Kingston, ON K7L 3N6, Canada

D. P. Munoz
Department of Psychology, Queen's University, Kingston, ON K7L 3N6, Canada

L. Itti
Neuroscience Program, University of Southern California, Los Angeles, CA 90089, USA

Introduction

Visual attention and eye movements enable us to interact with complex environments by selecting relevant information to be processed in the brain. To properly allocate attention, a network of brain resources is engaged, from low-level visual processing to motor control of gaze orienting [1]. This renders visual attention vulnerable to neurological disorders. Several neuropsychological and neuroimaging studies have demonstrated that damage in different areas of the attentional network can impair

distinct aspects of task performance or can reveal unusual patterns of brain activity in laboratory tasks that test for specific aspects of attention [2]. However, while in-depth clinical evaluation, structured behavioral tasks, and neuroimaging are extremely valuable and are the current gold standard for identifying particular impairments, they suffer from limitations that prevent their large-scale deployment: time and cost by limited numbers of medical experts, and inability of some patients (e.g., young children or some elderly) to either understand or comply with structured task instructions, or with the testing machinery or protocol.

Our core hypothesis is that natural attention and eye movement behavior—like a drop of saliva—contains a biometric signature of an individual and of her/his state of brain function or dysfunction. Such individual signatures, and especially potential biomarkers of particular neurological disorders, which they may contain, however, have not yet been successfully decoded. This is likely because of the high dimensionality and complexity of the natural stimulus (input space), of the stimulus to behavior transfer function (brain function), and of the behavioral repertoire itself (output space). We devised a simple paradigm that does not require expensive machinery, involves no preparation and no cognitive task for participants, is completed in 15 min, is portable for use outside large medical centers, and (after initial training of the machine learning algorithms) autonomously provides detailed decoding of an individual's signature.

We validated our technique with one neurodegenerative and two neurodevelopmental disorders that have been shown to involve deficits in visual attention and oculomotor functions. These deficits were exploited by our algorithm with features corresponding to oculomotor control, stimulus-driven (bottom-up) attention, and voluntary, contextual (top-down) attention. We first tested the algorithm on elderly participants with the neurodegenerative disorder, Parkinson's disease (PD) and validated the signature of PD discovered by our algorithm, because the behavioral deficits of PD are well understood. In short, PD is characterized by degeneration of dopaminergic neurons in the substantia nigra pars compacta, affecting basal ganglia processes, which subsequently impairs body movement (tremor, bradykinesia) and oculomotor movement (slower and shorter saccades) [3–5]. PD also impairs the prefrontal, premotor, motor, and basal ganglia networks [6], leading to deficits in attentional control; in particular, PD patients are less successful in inhibiting automatic saccades to a salient stimulus compared to controls [3, 4]. Therefore, we expected PD patients to show deficient oculomotor control, weakened top-down control, and stronger bottom-up guidance in natural viewing.

Next, we tested the algorithm on the two neurodevelopmental disorders at the other end of the age spectrum: attention deficit hyperactivity disorder (ADHD) and fetal alcohol spectrum disorder (FASD). Patients with ADHD or FASD demonstrate comparable deficits in visual attention tasks [7–11], but for different reasons. ADHD in childhood is characterized by delayed cortical maturation, dysfunction in dopamine transmission in the frontal cortex and/or basal ganglia [12], and decreased activity in frontal and striatal regions [13, 14]. These deficits result in difficulties in inhibiting premature responses (weakened top-down control), and thus patients appear more stimulus-driven (stronger bottom-up guidance) [7]. Oculomotor function seems relatively unimpaired, though previous studies have shown inconsistent findings [11]. On the other hand, FASD is caused by excessive maternal alcohol consumption, which results in malformation of the cerebral cortex, basal ganglia and cerebellum, and reduced overall brain and white-matter volumes [10, 15]. Deficits include impaired oculomotor functions [16], decreased top-down attentional control [17], and weakened bottom-up attention, possibly due to deficient visual sensory processing [18]. The weakened bottom-up guidance of children with FASD could be a differential factor between FASD and ADHD, because children with ADHD appear to be more stimulus-driven. For example, in pro-/anti-saccade tasks (where a pro-saccade requires participants to initiate an automatic eye movement to a visual stimulus, and an anti-saccade requires participants to make a voluntary eye movement in the opposite direction) [19], children with ADHD or FASD both made more directional errors in the anti-saccade task (implying difficulty in inhibiting automatic responses), but only children with FASD made more directional errors and had longer reaction time in the pro-saccade task (implying weakened stimulus-driven guidance) [7, 9]. While diagnosis of some subtypes of FASD is often assisted by the presence of dysmorphic facial features [20], the majority of affected children do not exhibit facial dysmorphology, and when these features are not obvious, there is a significant risk of misdiagnosis with ADHD [21]. Thus, the differential classification of ADHD versus FASD provides a difficult challenge for our method.

Methods

The experimental procedure is summarized in Fig. 1a. Participants' eye traces were recorded while watching 20 min of video. Participants were instructed to "watch and enjoy the clips." Five minutes of video was excluded from the analysis because of different lengths of the clip snippets for purposes beyond the scope of this study (see Supplementary Methods: Stimuli, Data Acquisition for detail). Each 30-s video clip was composed of 2–4-s clip snippets of unrelated scenes to minimize predictability and

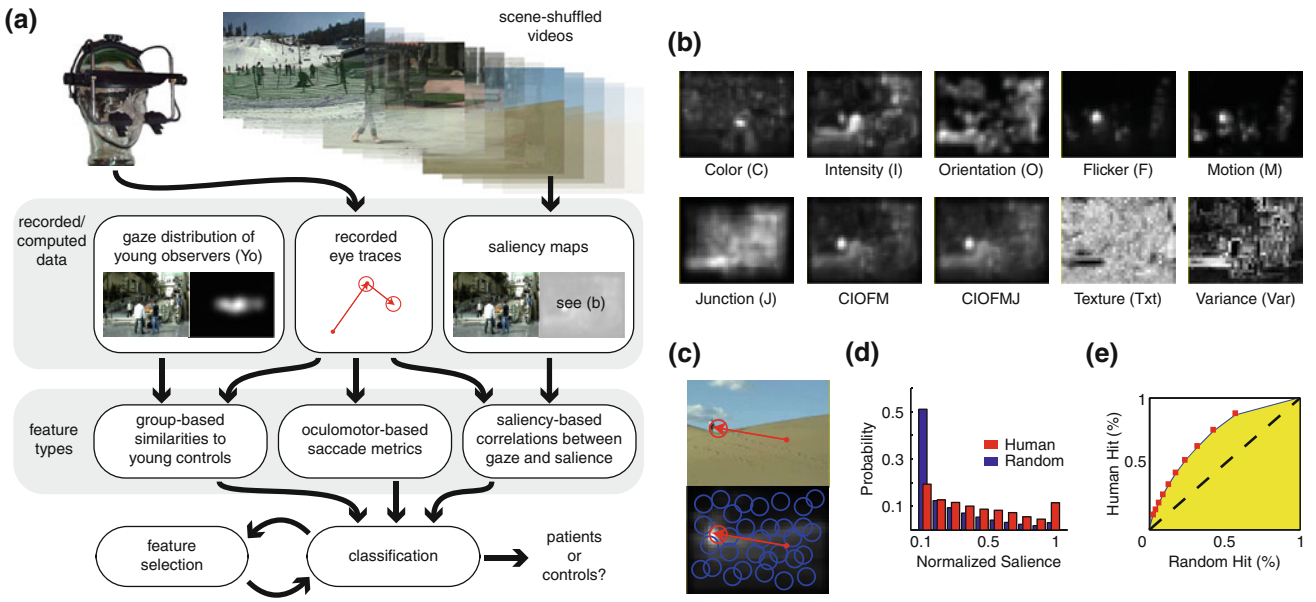

Fig. 1 Experimental and classification paradigms. **a** Participants freely viewed scene-shuffled videos (SV), and their eye movements were recorded. Saliency maps of each SV were computed using a computational model that mimics early visual processing. Next, we used the recorded eye movements to compute (1) oculomotor-based saccade metrics, (2) saliency-based correlations between saliency maps and gaze (bottom-up attention), and (3) group-based similarities in spatiotemporal distributions of gaze with reference to a database of control eye traces (top-down attention). These features were used in a classifier with a recursive feature selection method to identify important features that distinguished populations. **b** Ten saliency maps of different features (color, intensity, etc.) were computed, here illustrated for the video frame shown in (**a**) under "Saliency maps." Brighter shades of *grey* indicate stronger feature contrast at the corresponding image locations; for example, the *red* and *yellow* flowers between the two people elicit a strong response in the color contrast map. **c** To compute saliency-based or group-based features, each saliency map was sampled around the saccade target location (*red circle*) when a participant initiated a saccade (*red dot*). At the same time, 100 map values were randomly sampled from the map as a baseline (*blue circles*) for comparison. **d** Histograms were generated from both the human and random sample values. **e** Differences between human and random histograms were further summarized by ordinal dominance analysis to quantify the extent to which human observers gazed towards higher salience values than expected by chance in terms of the area under the curve (AUC, *yellow* region) (see Supplementary Methods: Computing Features for more detail)

to emphasize attentional deployment in new environments. Saliency maps (Fig. 1b; topographic maps that predict the locations of visually conspicuous stimuli based on low-level image properties; Supplementary Methods: Computing Saliency Maps from Stimuli) were computed for every frame [22], and correlations between model-predicted salience values and measured human saccade endpoints (gaze) were computed (Fig. 1c–e). Based on previous studies [3–18] of how the disorders may affect eye movement, we extracted a large number of features from the eye movement recordings (categorized into oculomotor-based, saliency-based, and group-based features; see Methods: Features) and built a classifier to differentiate patients and controls based on these features. We also analyzed the features for biomarkers through recursive evaluation, selection, and classification. Our workflow was inspired by successful application of advanced machine learning techniques to microarray analysis [23], here using similar techniques for the first time in high-throughput analysis of natural eye movement behavior.

Standard protocol approvals and patient consent

All experimental procedures were approved by the Human Research and Ethics Board at Queen's University, adhering to the guidelines of the Declaration of Helsinki and the Canadian Tri-Council Policy Statement on Ethical Conduct for Research Involving Humans.

Participants

This study describes data collected from 21 children with ADHD, 13 children with FASD, 14 elderly PD participants, 18 control children, 18 young controls, and 24 elderly controls (Table 1; Supplementary Methods: Participants, Diagnostic Criteria).

Features

From eye traces recorded while participants viewed short videos, we extracted three types of features that we

Table 1 Demographic data (see Supplementary Table S1 for full demographic data)

	n	Age (year)	Sub-type/severity	Medication
Ctrl. elderly	24	70.33 ± 7.53		
PD	14	67.43 ± 6.62	Hoehn and Yahr	Yes: 14
			Stage 2: 6	No: 0
			Stage 2.5: 6	
			Stage 3: 2	
Ctrl. young	18	23.17 ± 2.60		
Ctrl. child	18	10.67 ± 1.82		
ADHD	21	11.19 ± 1.83	Inattentive: 4	Yes: 16
			Hyperactive: 0	No: 5
			Combined: 19	
FASD	13	12.31 ± 2.10	FAS: 4	Yes: 10
			pFAS: 2	No: 3
			ARND: 7	

hypothesized would be differentially affected by disorders. First, oculomotor-based features were computed (e.g., distributions of saccade amplitudes and fixation durations) as they might reveal deficiencies in motor control of attention and gaze. Second, saliency-based features correlated participants' gaze to predictions from a computational model of visual salience [22], which has been previously shown to significantly predict which locations in a scene may more strongly attract attention of control subjects. We hypothesized that these features would reveal deficits in reflexive, stimulus-driven, or so-called "bottom-up" attention. The third type, group-based features, captured deviations in participants' gaze allocation onto our stimuli compared to a normative group of young adult controls. These features, we posited, might reveal impaired volitional, subject-dependent, or "top-down" attentional control, especially if differences were observed in group-based but not saliency-based features. Together, we utilized all these features to classify participants into clinical groups based on natural viewing behavior, the complexity of which imposed challenges in data analysis, but also revealed rich and profound information about the different populations.

The classifiers were built to discriminate patients from controls based on 15 core features from our three types: four oculomotor-based core features (distributions of saccade duration, inter-saccade interval, saccadic peak velocity, and saccade amplitude), ten saliency-based core features (differential distributions of salience values at human gaze vs other locations, using the ten saliency maps of Fig. 1b), and one group-based core feature (correlation between a patient's gaze and aggregate eye traces from a normative group of young adult controls, Fig. 1a). Each core feature was represented by several sub-features to capture the dynamics of free-viewing: each oculomotor-based core feature was subdivided into 12 sub-features [3 measures (lower quartile, medium, upper quartile) × 4 saccades (the 1st, 2nd, 3rd, and all saccades on each 2–4-s clip snippet) = 12 sub-features]; each saliency-based core feature was subdivided into 16 sub-features: 4 measures [area under the ROC curve (AUC; see Supplementary Methods: Computing Features) for low/medium/high salience bins] × 4 saccades, as was each group-based core feature: 4 measures (AUC, low/medium/high similarity bins) × 4 saccades. Thus, in total, 15 core features subdivided into 224 sub-features were used (Supplementary Table S2).

Classification and feature selection

Feature selection is a popular machine learning method to identify useful features and overcome situations where the number of features is possibly larger than the number of samples when training a classifier [24]. We performed feature selection with support vector machine-recursive feature elimination (SVM-RFE) [25], which has been used with great success in other fields (e.g., cancer classification with microarrays [25]). SVM-RFE consists of training a classifier and discarding the weakest feature iteratively until all features are eliminated. We used SVM-RFE to differentiate PD patients from elderly controls (binary classification), and multiple SVM-RFE (MSVM-RFE) [26] to distinguish children in the ADHD, FASD, and control groups (3-way classification). All classification accuracies reported were obtained using these two feature selection methods.

Performance of each classifier that used a particular selected subset of features was computed using 30 iterations of a repeated leave-one-out bootstrap validation [27]. This validation method was similar to the standard leave-one-out validation, which leaves one participant out for testing, but here the classifier was trained on the remaining participants that were bootstrapped (sample with replacement) ten times the number of these remaining participants. The performance was tested against permuted chance, which was the classification accuracy of a classifier trained on the same bootstrap structure but with randomly permuted class labels (class labels were randomly rearranged). Because classification accuracy varied with the number of features in the process of RFE, we tested the performance of classifiers by comparing the maximum accuracy obtained by the classifier trained with true labels to that obtained by the classifier trained with randomly permuted labels (permuted chance, the chance referred to in this article unless stated otherwise), regardless of how many features each classifier used to obtain maximum

accuracy (one-tail paired t test; Supplementary Methods: Classification and Feature Selection). All tests were Bonferroni corrected.

Results

Classifying PD and controls

Classification accuracy for 14 patients versus 24 age-matched controls reached 89.6 % (chance: 63.2 %, obtained by performing the same classification procedure with permuted class labels; Fig. 2a), with only 5 of 224 sub-features selected as most discriminative by the process of feature elimination (SVM-RFE). The confusion and sensitivity/specificity matrices reveal that the classifier made slightly more false negatives than false positives as we aimed to maximize overall classification performance. In scenarios where the classifier may be used for screening purposes, sensitivity of the classifier can be increased by assigning higher costs to missed PD patients and lower costs to false positives during training.

Our method not only differentiated PD from elderly controls [one-tail paired t test, $t(29) = 23.07$, $p < 0.01$], but also provided information about how PD affects eye movements, obtained by separately studying classification accuracy for oculomotor-based, saliency-based, or group-based features (Fig. 2b). PD patients demonstrated motor deficits as revealed by classification differences between them and controls in oculomotor features [considering only the 48 oculomotor-based sub-features, accuracy was 86.4 %, $t(29) = 28.02$, $p < 0.01$]. Oculomotor deficits have been attributed to dysfunction in the basal ganglia [28–30], crucial for voluntary saccade control [19]. Patient's top-down attention also differed from elderly controls [16 group-based sub-features, 74.6 %, $t(29) = 11.58$, $p < 0.01$], in agreement with previously reported impairment in voluntary attention, involving cortical and sub-cortical attention networks [28, 29, 31–33]. However, counter to our expectation that lower top-down control may give rise to higher reliance upon stimulus-driven salience, bottom-up attention of PD patients seemed unaffected, as saliency-based features showed no overall differences [160 saliency-based sub-features, 63.16 %, $t(29) = -4.10$, n.s]. It is possible that any higher reliance upon visually salient stimuli to guide gaze may have been offset by impaired salience computation because of deficient early visual processing in PD patients, as reported in previous laboratory studies [34] [see Supplementary Discussion of Neurological Implications: Parkinson's disease (PD) for more details relating the findings from previous studies to the results from classification].

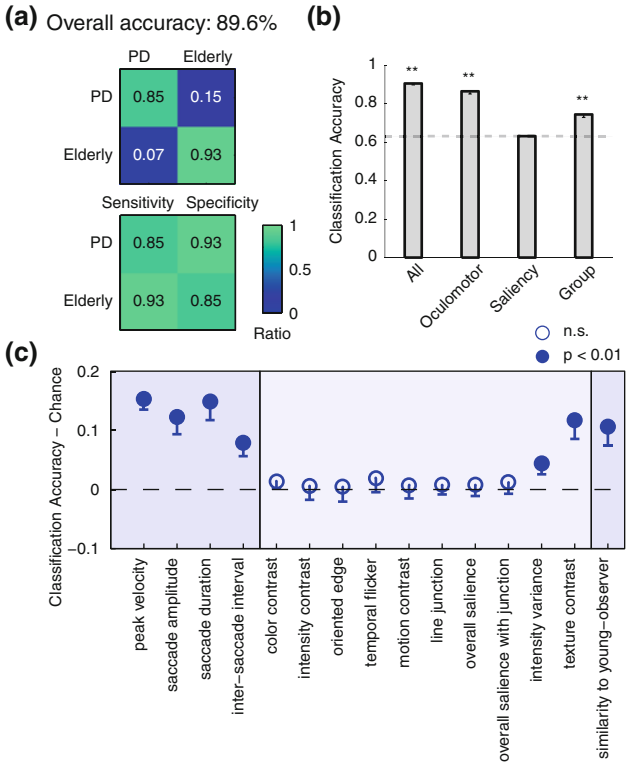

Fig. 2 Classification performance in differentiating PD patients from elderly controls at three granularities of starting feature sets: **a** all features, **b** the 3 feature types, and **c** the 15 core features (biometric signatures). **a** Starting with all 224 sub-features, PD patients were distinguished from elderly controls with 89.6 % accuracy after feature selection (SVM-RFE). Each row in the confusion matrix represents actual classes, and each column predicted classes. **b** PD and elderly controls differed significantly in oculomotor (starting with 48 sub-features) and group-based behavior (16 sub-features), but not in saliency processing (160 sub-features). *Asterisks* indicate cases where the classifiers performed significantly better than permuted chance (computed from training a classifier with randomly permuted class labels). *Dashed line* represents prior chance based on the number of controls and patients. **c** PD patients exhibited differences in saccade amplitude, duration, peak velocity, inter-saccade interval, intensity variance processing, texture saliency processing, and similarity to normative young observers. This pattern of differences yields the 15-component biometric signature of PD. *Dashed line* is the prior chance. Background colors separate oculomotor-based, saliency-based, and group-based features from left to right (error bars indicate 95 % confidence intervals after Bonferroni corrections). Significance level: $p < 0.01$, one-tailed paired t test ($df = 29$)

At a finer granularity, our method also permitted investigating whether each of our 15 core features was affected by PD. We tested 15 separate classifiers, each using only the 12 or 16 sub-features of a given core feature (with SVM-RFE). This yielded a 15-component biometric signature of PD (Fig. 2c). During natural viewing, PD patients demonstrated motor deficits as their saccades were of shorter amplitude and duration [classification accuracy: $t(29) > 9.62$, $p < 0.01$; direction of the effect: two-sample t test, $t(36) > 2.73$, $p < 0.01$]; peak velocity and inter-

saccade interval were also affected [$t(29) > 6.31$, $p < 0.01$], but without a unified upward or downward direction of effect among the 12 sub-features (Supplementary Methods: Direction of Effect). These observations are consistent with earlier structured-task studies, which showed shorter and slower voluntary saccades of PD patients toward pre-determined visual locations [3, 5, 28, 35], with less impairment for visually guided saccades [28, 35]. The classifier also found that PD and elderly controls differed in intensity variance [$t(29) = 4.96$, $p < 0.01$] and texture contrast [$t(29) = 8.36$, $p < 0.01$], though with mixed upward and downward effects among the involved sub-features, suggesting complex interactions between deficits that affect behavior in opposite directions: e.g., weakened top-down control (stronger bottom-up) and impaired saliency computation (weaker bottom-up). Deficits in voluntary control and top-down attention were also revealed by different similarities to our normative young observers between PD patients and elderly controls [$t(29) = 7.06$, $p < 0.01$].

Classifying ADHD, FASD, and control children

Classification accuracy with MSVM-RFE for 21 children with ADHD vs. 13 children with FASD vs. 18 control children reached 77.3 % (chance 40.4 %) with 19 of all 224 sub-features (Fig. 3a). With these 19 features, the average two-way classification accuracy for ADHD versus control was 83.3 % (chance 53.8 %); FASD versus control was 79.2 % (chance 58.1 %); ADHD versus FASD was 90.4 % (chance 61.8 %). Rates of miss and false alarm errors were balanced, except for a slightly higher miss rate for FASD, as the classifier aimed to maximize overall accuracy.

Our method further examined which of the three feature types contained differential information among the three groups of children (Fig. 3b). Classification accuracies were significantly above chance with the saliency-based [50.8 %, $t(29) = 4.04$, $p < 0.05$], but not with the oculomotor-based features [40.5 %, $t(29) = -5.28$, n.s] and the group-based features [45.7 %, $t(29) = 1.03$, n.s.]. When comparing each pair of the three child groups, first, children with ADHD and controls were distinguished significantly in saliency-based features [78.2 %, $t(29) = 12.68$, $p < 0.01$]; second, children with FASD and controls differed in both saliency-based features [77.6 %, $t(29) = 9.95$, $p < 0.01$] and group-based features [69.8 %, $t(29) = 6.01$, $p < 0.01$]; lastly, children with ADHD and FASD showed no differentiability by each feature type alone, but they could be distinguished with all feature types together [$t(29) < 22.96$, $p < 0.01$]. Although we focus on classification performance, these results are in line with earlier studies that showed how children with ADHD have difficulties in inhibiting premature responses and thus appear more stimulus-driven [7], as well as studies that demonstrated how children with FASD have atypical top-down [8, 9, 17] and bottom-up [18] attentional control (see Supplementary Discussion of Neurological Implications: ADHD, FASD, and ADHD versus FASD for more details pertaining to previous studies and the present results). However, when we examined whether the saliency-based and group-based sub-features showed larger feature values in one population than in the other, we found mixed directions of effect among the sub-features of both feature types, indicating that the disorder impacts natural viewing behavior in more than one single unified manner (e.g., impaired response inhibition [7, 9], but also possibly weakened early visual processing [36–38]). The quantitative predictions of our classifier for every sub-feature provide for the first time a rich basis to further investigate these complex effects from a neurological viewpoint.

At the level of the 15 core features, our method yielded clearly distinct biometric signatures for ADHD versus FASD (Fig. 3c), thus successfully teasing apart the two disorders along 15 important dimensions. For children with ADHD, the best feature differentiating them from control children was texture processing [$t(29) = 15.67$, $p < 0.01$]; children with ADHD showed a higher correlation with texture contrast [two-sample t test, $t(37) = 2.75$, $p < 0.01$; Fig. 3c], in line with previously reported tactile texture sensitivity [39–41]. Thus, the current results suggest this may not be limited to the tactile domain. Propensity to look toward color contrast [36, 37] [$t(29) = 5.63$, $p < 0.01$] and oriented edges [$t(29) = 6.72$, $p < 0.01$] was also discriminative between children with ADHD and controls. Oriented edges are important to perceptually construct the contour and shape of objects. For children with FASD, line junctions, overall salience, and texture contrast were discriminative [$t(29) > 4.92$, $p < 0.01$]. To our knowledge, no previous study has investigated how ADHD might affect processing of oriented edges, nor how different domains of salient features may be affected by FASD. The discovery of these features by our classifier thus suggests interesting new research directions.

Sub-features selected by the SVM-RFE process

Finally, we investigated which collections of sub-features best differentiated the populations based on the result of feature selection (SVM-RFE). The top five sub-features that classified PD from elderly controls and their normalized feature values are shown in Fig. 4a. The feature selection method found a collection of five oculomotor sub-features that reliably differentiate PD from elderly controls. On the other hand, the top 19 features for differentiating each pair of the three child populations

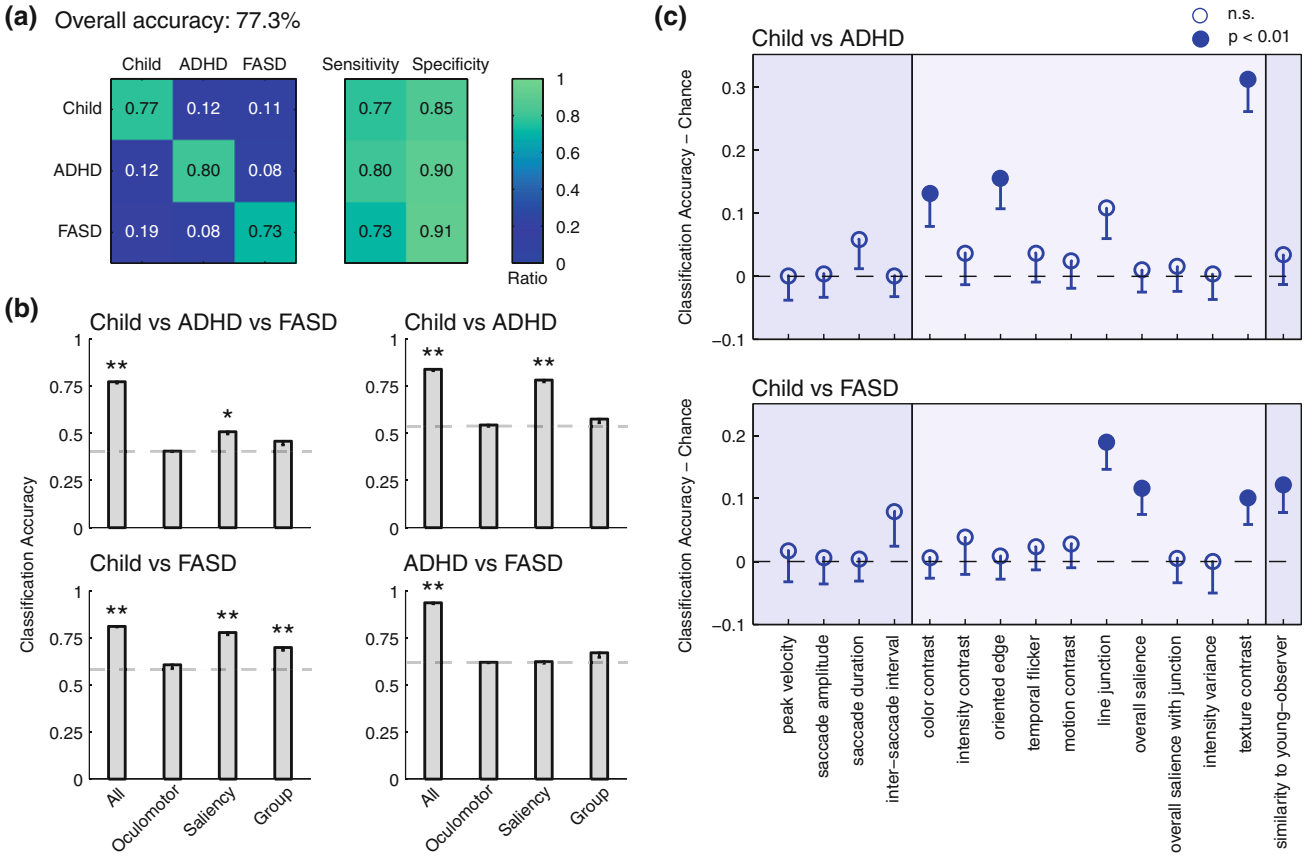

Fig. 3 Classification performance for children with ADHD, FASD, and control children for: **a** all features, **b** the 3 feature types, and **c** the 15 core features (biometric signatures). **a** Starting with all sub-features, children with ADHD, FASD, and control children were best classified with 77.3 % accuracy (ADHD: sensitivity 80 %, specificity 90 %; FASD: sensitivity 73 %, specificity 91 %) after feature selection (MSVM-RFE). Format is as in Fig. 2. **b** Classifying the three child groups with different feature sets demonstrated that they differed significantly in saliency-based behavior (upper-left sub-plot). Children with ADHD differed from control children in saliency-based features, whereas children with FASD differed from controls in both saliency-based and group-based features, and children with ADHD and those with FASD could only be distinguished with all three feature types together. **c** The 15-component biometric signature of ADHD and FASD. Children with ADHD compared to control children demonstrated significantly different sensitivity in color contrast and oriented edges, as well as increased sensitivity to texture contrast. Children with FASD, in contrast, showed a different signature that involved differences in similarity to young observers in gaze distribution, sensitivity to line junctions, and sensitivity to overall salience, as well as increased sensitivity to texture contrast. Background colors separate oculomotor-based, saliency-based, and group-based features from left to right (see Fig. 2 for the computation of chance level, error bars, statistical tests, and significance level)

(ADHD, FASD and control) spanned all three broad feature types (Fig. 4b). While some core features of the selected sub-features failed to differentiate the populations when considered in isolation, they are important complementary features for the classifiers to separate the groups. Obviously the pattern of feature values observed here is complex, indicating that sophisticated classifiers were indeed necessary to discover the subsets of features that yielded the best classification accuracy. To visualize how well our approach was able to cluster individuals into separate groups, we further reduced the dimensionality of our results using Linear Discriminant Analysis, which finds axes that best separate each pair of groups (Fig. 4c, d). This analysis allowed us to validate our method by demonstrating clearly distinct clusters based on the features selected by our classifiers. We suggest that similar clustering techniques could be employed in future studies of other disorders and to possibly discover different sub-population clusters within patient groups that were previously considered homogeneous per standard medical assessment.

Discussion

This study revealed different biometric profiles of oculomotor function and attention allocation among PD, ADHD, and FASD patient groups through quantitative analysis of natural viewing eye traces. Our automated SVM-RFE process discovered that PD patients were best

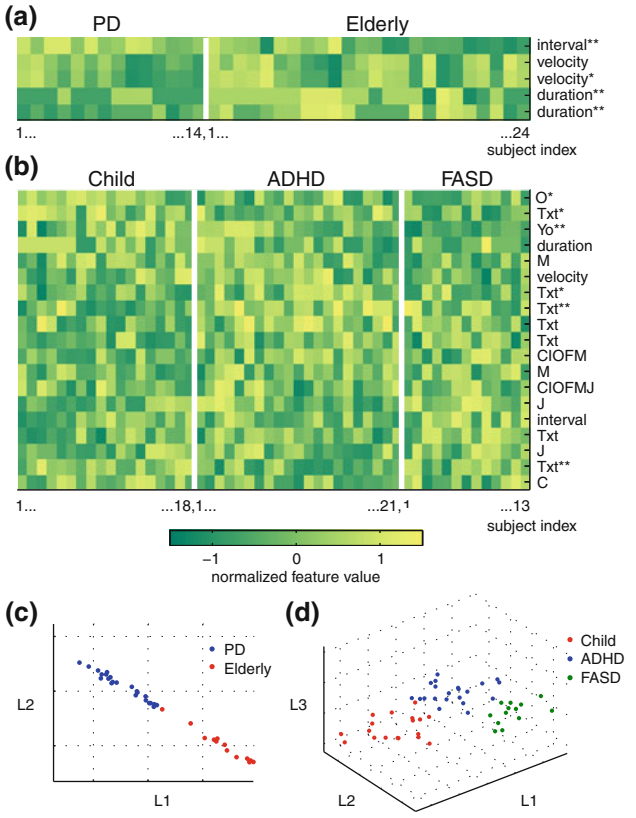

Fig. 4 Sub-features selected by SVM-RFE. **a** Normalized values for the top five ranked sub-features selected by SVM-RFE for PD. Sub-feature's names were replaced by their corresponding core features for simplicity (see Supplementary Fig. S1 for sub-feature names; e.g., two different sub-features of the velocity core feature were selected). Feature values are standardized z-scores filtered by an arctangent function. *Rows* represent the top five ranked sub-features. *Columns* represent 38 participants, and the white vertical line separates the two populations. **b** Normalized feature values for the top 19 ranked sub-features selected by MSVM-RFE that best classified children with ADHD, FASD, and control children. Note that most of the sub-features discovered by the classifiers belonged to the saliency-based feature type. Features and participants were re-arranged so that high feature values were better clustered at the diagonal of the plot. **c** 14 PD and 24 elderly controls were separated into two different clusters as revealed by linear discriminant analysis (LDA), which finds the dimensions (L1 and L2) from the top five sub-features in (**a**) that best distinguished the two groups. **d** Similarly, LDA found the three dimensions (L1, L2, and L3) from the top 19 sub-features in (**b**) that best differentiated every pair among 21 children with ADHD, 13 children with FASD, and 18 control children. The three child groups are clearly separated in these dimensions, even though clusters in (**b**) are less visually distinct. (ANOVA $^{**}p < 0.01$; ANOVA $^{*}p < 0.05$)

discriminated from elderly controls by oculomotor-based features, implying that motor deficits are more apparent than attention deficits for PD patients during free viewing. In contrast, children with ADHD or FASD were best distinguished from controls by saliency-based features, suggesting that the disorders affect their bottom-up attention. The disorders also influence overall attention allocation in every patient group, as group-based features showed differentiability for clinical and control populations (see Supplementary Discussion for our interpretations of the particular features identified by our method and the corresponding neurological implications in each disorder). By identifying features that are most discriminative among populations, our technique provides new insights into the nature of the different disorders and their interactions with attentional control. The encouraging results obtained here with diseases that lie on both ends of the age spectrum suggest that the proposed approach may generalize to additional disorders that affect attention and oculomotor systems. The fact that our paradigm alleviates the need for structured tasks is of great importance because the approach can be applied to a wider range of populations, including very young children who cannot understand the instructions of experiments or individuals who have cognitive impairment.

Our method robustly differentiates disorders that may have overlapping behavioral phenotypes (ADHD and FASD) but that nonetheless affect visual processing differently. Overall, we suggest that with natural scene videos, participants' natural viewing behaviors are evoked, and their eye movement patterns contain unique and revealing information about their cognitive and motor processes. One of the strengths of this study is that it is a general framework that could identify such information in several patient populations. In the future, with better understanding of differences in cognitive control, attention, and oculomotor systems of patients with these disorders, the experiment could be further shortened by selecting stimuli that maximally evoke different eye movement patterns between populations. This would also provide for a better understanding of novel behavioral differences that were revealed by this study, such as the discovery of edge processing differences in children with ADHD (see Supplementary Discussion: Future Directions and Study Limitations). In summary, our method provides for the first time an objective, automated, high-throughput, time- and cost-effective tool that can screen large populations and that, through clustering, may further discover new disease subtypes and assist making more precise medical diagnoses. Future benefits of our method may include earlier and more accurate identification of neurological disorders and subtypes.

Acknowledgments We thank the National Science Foundation (CRCNS grant number BCS-0827764), the Army Research Office (grant nos. W911NF-08-1-0360 and W911NF-11-1-0046), the Human Frontier Science Program (grant RGP0039/2005-C), and the Canadian Institutes of Health Research (grant no. ELA 80227) for supporting this study. IGMC was supported by a scholarship from the Canadian Institutes for Health Research, and DPM was supported by the Canada Research Chair program.

Conflicts of interest None.

References

1. Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215
2. Corbetta M, Kincade MJ, Lewis C, Snyder AZ, Sapir A (2005) Neural basis and recovery of spatial attention deficits in spatial neglect. Nat Neurosci 8:1603–1610
3. Chan F, Armstrong IT, Pari G, Riopelle RJ, Munoz DP (2005) Deficits in saccadic eye-movement control in Parkinson's disease. Neuropsychologia 43:784–796
4. Cameron IGM, Watanabe M, Pari G, Munoz DP (2010) Executive impairment in Parkinson's disease: response automaticity and task switching. Neuropsychologia 48:1948–1957
5. Terao Y, Fukuda H, Yugeta A, Hikosaka O, Nomura Y, Segawa M et al (2011) Initiation and inhibitory control of saccades with the progression of Parkinson's disease—changes in three major drives converging on the superior colliculus. Neuropsychologia 49:1794–1806
6. Wise SP, Murray EA, Gerfen CR (1996) The frontal cortex-basal ganglia system in primates. Crit Rev Neurobiol 10:317–356
7. Munoz DP, Armstrong IT, Hampton KA, Moore KD (2003) Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. J Neurophysiol 90:503–514
8. Green CR, Mihic AM, Brien DC, Armstrong IT, Nikkel SM, Stade BC et al (2009) Oculomotor control in children with fetal alcohol spectrum disorders assessed using a mobile eye-tracking laboratory. Eur J Neurosci 29:1302–1309
9. Green CR, Munoz DP, Nikkel SM, Reynolds JN (2007) Deficits in eye movement control in children with fetal alcohol spectrum disorders. Alcohol Clin Exp Res 31:500–511
10. Kodituwakku PW (2007) Defining the behavioral phenotype in children with fetal alcohol spectrum disorders: a review. Neurosci Biobehav Rev 31:192–201
11. Karatekin C (2007) Eye tracking studies of normative and atypical development. Dev Rev 27:283–348
12. Vaidya CJ, Austin G, Kirkorian G, Ridlehuber HW, Desmond JE, Glover GH et al (2011) Selective effects of methylphenidate in attention deficit hyperactivity disorder: a functional magnetic resonance study. Proc Natl Acad Sci USA 95:14494–14499
13. Amen DG, Carmichael BD (1997) High-resolution brain SPECT imaging in ADHD. Ann Clin Psychiatry 9:81–86
14. Zametkin AJ, Nordahl TE, Gross M, King AC, Semple WE, Rumsey J et al (1990) Cerebral glucose metabolism in adults with hyperactivity of childhood onset. New Engl J Med 15(323): 1361–1366
15. Riley EP, McGee CL (2005) Fetal alcohol spectrum disorders: an overview with emphasis on changes in brain and behavior. Exp Biol Med 230:357–365
16. Kalberg WO, Provost B, Tollison SJ, Tabachnick BG, Robinson LK, Eugene Hoyme H et al (2006) Comparison of motor delays in young children with fetal alcohol syndrome to those with prenatal alcohol exposure and with no prenatal alcohol exposure. Alcohol Clin Exp Res 30:2037–2045
17. Lezak MD (1995) Neuropsychological assessment, 3 edn. Oxford University Press, Oxford
18. Carr JL, Agnihotri S, Keightley M (2010) Sensory processing and adaptive behavior deficits of children across the fetal alcohol spectrum disorder continuum. Alcohol Clin Exp Res 34:1022–1032
19. Munoz DP, Everling S (2004) Look away: the anti-saccade task and the voluntary control of eye movement. Nat Rev Neurosci 5:218–228
20. Jones KL, Smith DW (1973) Recognition of the fetal alcohol syndrome in early infancy. Lancet 3(302):999–1001
21. Crocker N, Vaurio L, Riley EP, Mattson SN (2009) Comparison of adaptive behavior in children with heavy prenatal alcohol exposure or attention-deficit/hyperactivity disorder. Alcohol Clin Exp Res 33:2015–2023
22. Itti L, Koch C (2001) Computational modelling of visual attention. Nat Rev Neurosci 2:194–203
23. Golub TRR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP et al (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 15(286):531–537
24. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3:1157–1182
25. Guyon I, Weston J, Barnhill S, Vapnik V (2002) Gene selection for cancer classification using support vector machines. Mach Learn 46:389–422
26. Zhou X, Tuck DP (2007) MSVM-RFE: extensions of SVM-RFE for multiclass gene selection on DNA microarray data. Bioinformatics 23:1106–1114
27. Jiang W, Simon R (2007) A comparison of bootstrap methods and an adjusted bootstrap approach for estimating the prediction error in microarray classification. Stat Med 26:5320–5334
28. Briand KA, Strallow D, Hening W, Poizner H, Sereno AB (1999) Control of voluntary and reflexive saccades in Parkinson's disease. Exp Brain Res 129:38–48
29. Fukushima J, Fukushima K, Miyasaka K, Yamashita I (1994) Voluntary control of saccadic eye movement in patients with frontal cortical lesions and parkinsonian patients in comparison with that in schizophrenics. Biol Psychiatry 36:21–30
30. Watanabe M, Munoz DP (2011) Probing basal ganglia functions by saccade eye movements. Eur J Neurosci 33:2070–2090
31. Kori A, Miyashita N, Kato M, Hikosaka O, Usui S, Matsumura M (1995) Eye movements in monkeys with local dopamine depletion in the caudate nucleus. II: deficits in voluntary saccades. J Neurosci 15:928–941
32. Pierrot-Deseilligny C, Milea D (2004) Eye movement control by the cerebral cortex. Curr Opin Neurol 17:17–25
33. Leh SE, Petrides M, Strafella AP (2010) The neural circuitry of executive functions in healthy subjects and Parkinson's disease. Neuropsychopharmacology 35:70–85
34. Bodis-Wollner I (2003) Neuropsychological and perceptual defects in Parkinson's disease. Parkinsonism Relat Disord 9(2):S83–S89
35. White OB, Saint-Cyr JA, Tomlinson RD, Sharpe JA (1983) Ocular motor deficits in Parkinson's disease. II: control of the saccadic and smooth pursuit systems. Brain 106 (Pt 3):571–587
36. van der Stelt O, van der Molen M, Boudewijn Gunning W, Kok A (2001) Neuroelectrical signs of selective attention to color in boys with attention-deficit hyperactivity disorder. Cognitive Brain Res 12:245–264
37. Jonkman LM, Kenemans JL, Kemner C, Verbaten MN, van Engeland H (2004) Dipole source localization of event-related brain activity indicative of an early visual selective attention deficit in ADHD children. Clin Neurophysiol 115:1537–1549
38. Burden MJ, Westerlund A, Muckle G, Dodge N, Dewailly E, Nelson CA et al (2011) The effects of maternal binge drinking during pregnancy on neural correlates of response inhibition and memory in childhood. Alcohol Clin Exp Res 35:69–82
39. Ermer J, Dunn W (1998) The sensory profile: a discriminant analysis of children with and without disabilities. Am J Occup Ther 52:283–290

40. Mangeot SD, Miller LJ, McIntosh DN, McGrath-Clarke J, Simon J, Hagerman RJ et al (2001) Sensory modulation dysfunction in children with attention-deficit–hyperactivity disorder. Dev Med Child Neurol 12(43):399

41. Parush S, Sohmer H, Steinberg A, Kaitz M (2011) Somatosensory functioning in children with attention deficit hyperactivity disorder. Dev Med Child Neurol 39:464–468