Unravelling Neurodivergent Gaze Behaviour through Visual Attention Causal Graphs

Giuseppe Cartella[1], Vittorio Cuculo[1], Alessandro D'Amelio[2], Rita Cucchiara[1], and Giuseppe Boccignone[2]

[1] University of Modena and Reggio Emilia, Italy
{name.surname}@unimore.it

[2] University of Milan, Italy
{name.surname}@unimi.it

Abstract. Can the very fabric of how we visually explore the world hold the key to distinguishing individuals with Autism Spectrum Disorder (ASD)? While eye tracking has long promised quantifiable insights into neurodevelopmental conditions, the causal underpinnings of gaze behaviour remain largely uncharted territory. Moving beyond traditional descriptive metrics of gaze, this study employs cutting-edge causal discovery methods to reconstruct the directed networks that govern the flow of attention across natural scenes. Given the well-documented atypical patterns of visual attention in ASD, particularly regarding socially relevant cues, our central hypothesis is that individuals with ASD exhibit distinct causal signatures in their gaze patterns, significantly different from those of typically developing controls. To our knowledge, this is the first study to explore the diagnostic potential of causal modeling of eye movements in uncovering the cognitive phenotypes of ASD and offers a novel window into the neurocognitive alterations characteristic of the disorder.

Keywords: Causal Graphs · Eye Tracking · Autism Spectrum Disorder · Visual Attention · Causal Discovery

1 Introduction

Decades of meticulous inquiry into the nature of eye movements and visual attention have shaped the delicate interplay between the determinism and randomness of where our eyes will alight as arising from the insistent call of bottom-up cues (salient features due to stark contrasts of luminance, colour and motion, e.g. [16]) that harmonizes with the deliberate direction of top-down influences [6,9,16]. In this realm, attention transcends mere reaction, becoming a nuanced expression of our inner world, shaped by the demands of our current endeavours, the rich tapestry of our accumulated knowledge, and the anticipation of what might lie ahead [9, 27, 33]. Even further, it has been argued that, under the assumption of perception as hypothesis testing, eye movements can be considered as experiments, in which data are gathered to test (counterfactual) hypotheses or beliefs about how those data are caused [11, 13].

Specifically, the idea that attention allocation follows a causal structure suggests that top-down attention control may be represented as a form of high-level causal reasoning that guides the observer attentive dynamics towards specific items. Eventually, the sequence of fixations chosen to sample the scene may be influenced by underlying cognitive factors such as semantic relatedness [3], task relevance, social meaning [1], and individual cognitive characteristics [26]. In fact, humans possess sophisticated mechanisms that enable reasoning about cause and effect in both abstract and formal terms [21]. Interestingly enough, recent research has demonstrated how these mechanisms, particularly counterfactual reasoning, can be revealed by an observer's gaze behaviour [13]. This causal influence likely reflects a dynamic interaction between perceptual inputs and the observer's internal state, suggesting that visual behaviour can serve as an indirect window into higher-level cognitive functioning.

This perspective may have profound implications for developmental and clinical research. Autism Spectrum Disorder (ASD) is a complex disorder characterised by dysfunctions of between- and within-networks functional connectivity involving the somatomotor, visual, salience and default mode networks (DMN) [31]. The DMN critically modulates the theory of mind (ToM), social interaction, emotions, etc. Notably, individuals with ASD exhibit atypical patterns of visual attention that are often characterized by diminished attention to socially salient stimuli, such as faces or eye regions [22, 25], though variations might occur as a function of the social context in which stimuli are embedded. Traditional metrics of gaze behaviour, such as total fixation duration or number of fixations on regions of interest, even when complemented by sophisticated statistical analyses (e.g., [15]), provide valuable descriptive information but fail to capture the underlying causal structure of attentional deployment (cfr., Section 2). Instead, modelling gaze as a phenomenon governed by causal relationships potentially allows for a deeper understanding of how attention is orchestrated moment-to-moment, and how this differs between groups of individuals.

The chief concern of the present work is to verify whether the causal structure governing the dynamics of attention allocation on natural scenes (images) has the potential to differentiate subjects with Autism Spectrum Disorder (ASD) from typically developing (TD) controls. Although it is largely agreed that eye tracking offers measurable variables that are promising in the study of neurodevelopmental disorders [12, 29] and could serve as biomarkers for early diagnosis, monitoring and further care interventions [12], the very causal nature of gaze unfolding is substantially unexplored in this field.

Here, more specifically, by means of a state-of-the-art causal discovery method from time series data [24], we estimate the directed acyclic graphs (DAGs) describing the causal mechanisms of attention allocation from eye-tracking data (Section 3). Crucially, these causal graphs capture not only the statistical patterns of gaze transitions on specific regions of interest but also infer the directional causal influence that fixating on a given item has on another, shedding light on the hierarchical structure of attentional control. We hypothesize that individuals with ASD will exhibit distinct causal structures of gaze deployment,

reflecting alterations in how cognitive factors modulate attention. Sections 4 and 5 report and discuss results achieved on a publicly available dataset.

To the best of our knowledge, this is the first study to investigate the diagnostic potential of causal modelling of gaze behaviour to discover the cognitive phenotypes of autism spectrum disorder.

2 Related Works

Modelling gaze behaviour has become an important area of study in computer vision and cognitive science [2], particularly for understanding attention mechanisms and supporting diagnostic applications in neurodevelopmental disorders.

Gaze behaviour has proven especially valuable in the context of ASD detection and ASD/ADHD comorbidity [15]. Some studies have developed frameworks that utilize multi-modal behavioural data, including eye-tracking, to distinguish ASD from TD [4, 17, 19]. Others have proposed models capable of simulating the dynamic gaze behaviour (scanpaths) of neurodivergent individuals [5]. Additionally, multi-level saliency models have been introduced to capture distinct attentional patterns in ASD, such as a central fixation bias and reduced attention to faces and socially relevant cues [30]. The Saliency4ASD Grand Challenge further advanced this field by providing standardized datasets and tools to support the development of visual attention models tailored to ASD characteristics [14].

From a biometrics angle, eye movements have been modelled using stochastic processes rooted in foraging theory, offering personalized behavioural signatures for identification [10]. Other studies have analysed gaze in social contexts, linking it to personality traits using probabilistic models [7]. Similarly, sequence models trained on eye movement data have shown success in ADHD detection, emphasizing the importance of dynamic gaze features and task design [8].

Notably, recent solutions have started to employ high-level graph representations to model human behaviour by means of eye-tracking data. One recent approach is the Attention Graph, which encodes saliency and semantic-driven scanpaths in a graph-based structure to capture common gaze behaviours across individuals [32]. In a different vein, GazeGraph introduced a flexible and generalizable framework that combines spatial-temporal gaze graphs with deep learning and meta-learning for quick adaptation and robust context recognition [18].

While such representations enhance both interpretability and utility in tasks such as ASD screening and age classification [32], their modelling approach is limited to quantifying statistical associations, without accounting for higher-level causal drivers of attention allocation. Notably, as mentioned in Section 1, recent research has shown that counterfactual simulation during gaze tasks plays a significant role in causal reasoning [13], supporting the development of causal frameworks for attention modelling. This work takes a first step in that direction by proposing Visual Attention Causal Graphs, aiming to integrate attention modelling and causal reasoning for a deeper understanding of gaze behaviour.

3 Proposed Method

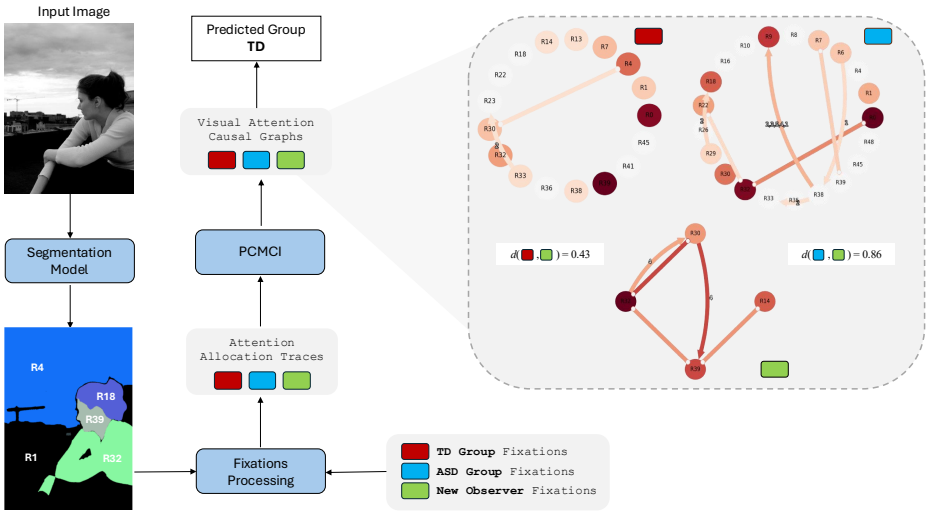


Fig. 1: The proposed method at a glance. Semantically relevant items are automatically extracted from the input image using a segmentation model. Visual Attention Causal Graphs (VA-CGs) are then generated as templates for the ASD (blue) and TD (red) observer groups. A VA-CG is also constructed for the new observer (green). Finally, the new observer is classified as either ASD or TD by minimizing the Jaccard distance between their VA-CG and the template VA-CGs.

An overview of our approach is summarised in Fig. 1. A causal graph is estimated, which describes the mechanisms governing the dynamics of attention allocation towards semantically relevant *items* within a scene. The goal is to validate the assumption that such graphs, referred to as Visual Attention Causal Graphs (VA-CG), effectively capture the observer's top-down attention allocation process and provide insights into their behavioral cues. Specifically, we aim to quantify their ability to assess high-level cognitive functioning, with a focus on quantitative screening for autism spectrum disorder (ASD). The approach is divided into three phases:

1. *Construction of Aggregated Item-Level Attention Allocation Traces:* The visual attention behavior of a group of subjects is represented as a multivariate time-series that describes the aggregated group-level attention toward each semantically relevant item inside the scene.
2. *Estimation of the Visual Attention Causal Graph (VA-CG):* The causal mechanism of attention allocation dynamics is reformulated as a multivariate time-series causal discovery problem. The state-of-the-art PCMCI causal discovery method [24] is applied to estimate the associated causal graph.

3. *ASD vs. TD Classification via the Obtained VA-CG Representations:* The final phase involves using the VA-CG representations to classify subjects as either exhibiting traits of autism spectrum disorder (ASD) or as typically developing (TD).

3.1 Aggregated Item-Level Attention Allocation Traces

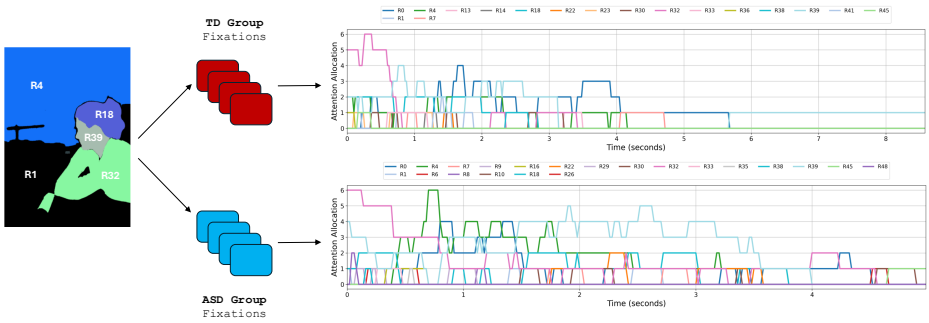


Fig. 2: Illustration of the construction of multivariate time-series representing group-level visual attention allocation on a scene. The input image is segmented into semantically meaningful regions (e.g., R1, R4, R18), and fixation data from the TD (top) and ASD (bottom) observer groups are aggregated over time. The resulting time-series plots show the temporal allocation of visual attention to each region, capturing group-level dynamics in visual attention distribution.

The first step of the proposed approach implies the construction of the multivariate time-series that describes the aggregated group-level attention toward semantically relevant items. Consider a stimulus j (image), observed by a group of S subjects with their eye movements tracked in real-time.

A segmentation approach is adopted to define semantically relevant items within an image, where each item corresponds to a distinct region or object [20]. For each time instant, we quantify the visual attention allocated to each item by counting the number of fixations directed towards it by the subjects.

This procedure generates a multivariate time series, where each time series corresponds to an individual item, describing the aggregated attention allocation across multiple subjects over time. The resulting data, shown in Fig. 2, captures the group-level temporal dynamics of visual attention at the item level, providing insight into how attention is distributed across different regions of the image as it is processed by the viewers.

3.2 Visual Attention Causal Graphs (VA-CG)

To estimate the causal graph from the obtained aggregated item-level attention allocation traces, we employ the PCMCI causal discovery method [24].

PCMCI (Peter–Clark Momentary Conditional Independence) is a method designed to uncover causal relationships within multivariate time series data. By analyzing the temporal dependencies among variables, PCMCI infers causal structures which can be compactly represented using Directed Acyclic Graphs (DAGs). As with any causal discovery framework, PCMCI relies on a set of foundational assumptions. These include the causal Markov condition and the faithfulness assumption [28], as well as specific conditions tailored to time series: causal stationarity (the assumption that causal relationships remain constant over time), the absence of instantaneous (contemporaneous) causal links, and causal sufficiency - meaning that all causally relevant variables are observed and included in the model [24].

In this study, we apply PCMCI to the aggregated semantic-level attention allocation traces described above. Specifically, for each time series $x_i(t)$, corresponding to the amount of attention allocated to a given item at time t, PCMCI employs Conditional Independence (CI) testing to detect causal dependencies. Causal links between traces $x_i(t)$ and $x_j(t)$ are determined by evaluating their conditional independence ($\perp$) with respect to a set of potentially mediating variables $\mathbf{z}(t)$:

$$x_i(t) \perp x_j(t) \,|\, \mathbf{z}(t),$$

where $\mathbf{z}(t)$ encompasses other variables that may influence the interaction between $x_i(t)$ and $x_j(t)$.

The application of PCMCI requires the specification of several key hyperparameters:

- *Conditional Independence Test:* The choice of CI test is critical, with available options including Partial Correlation (ParCorr), its robust variant (r-ParCorr), as well as discrete-variable tests such as the G-squared test. Selection depends on the nature of the data, its distributional properties and sample size.
- *Time Lag Range (τ_{min}, τ_{max}):* Defines the minimum and maximum temporal delays over which causal relationships are assessed. Larger lag windows facilitate the discovery of long-range dependencies but at the cost of increased computational complexity.
- *Significance Threshold (α):* Specifies the alpha level for the CI tests, serving to filter out statistically insignificant or spurious causal links.

By analysing directed, time-lagged dependencies among the attention traces, PCMCI constructs a causal graph that reflects the structure of attention flow across the items in a scene. We refer to such graphs as Visual Attention Causal Graphs (VA-CG). Time-unrolled versions of VA-CGs, estimated on the same image for both TD and ASD groups, are shown in Fig. 3.

3.3 ASD Detection via VA-CGs

For each image j in the dataset, available subjects are used to build the Aggregated Item-Level Attention Allocation Traces for both the ASD and TD groups.

These traces are then used to construct the VA-CGs for the ASD (G_j^{ASD}) and TD (G_j^{TD}) groups. These graphs serve as templates for classifying a new subject's (s) scanpath on the j-th image by computing its visual attention causal graph G_j^s using the procedure described in Section 3.

The assignment of the s-th subject's eye tracking data to one of the groups is determined by calculating the Jaccard distance between the ASD and TD VA-CGs and the new subject's causal graph:

$$\text{group}_j^* = \arg \min_{g \in \{\text{ASD,TD}\}} d\left(G_j^g, G_j^s\right).$$

Specifically, given a threshold $\beta \in (0, \alpha)$, that sets the minimum causal strength, the set of links (edges) considered in the graph G_j with N nodes is defined as:

$$\mathcal{E}_j = \{(m, n, \tau) \mid P_{m,n,\tau} < \beta \ \wedge \ |T_{m,n,\tau}| > 0\}$$

where $T \in \mathbb{R}^{N \times N \times (\tau_{\max}+1)}$ and $P \in [0,1]^{N \times N \times (\tau_{\max}+1)}$ are the matrices of test statistic values and the matrices of p-values for the graph G_j, respectively.

The distance between the two graphs, based on the Jaccard distance, is:

$$d(G_j^g, G_j^s) = 1 - \frac{|\mathcal{E}_j^g \cap \mathcal{E}_j^s|}{|\mathcal{E}_j^g \cup \mathcal{E}_j^s|}.$$

Following previous work [4,32], more robust subject-level classification can be obtained by a majority voting scheme based on the predictions across multiple stimuli.

4 Experimental analysis

4.1 Dataset

We adopt the Saliency4ASD dataset [14], which offers a collection of 300 images accompanied by eye-tracking records from children diagnosed with ASD and TD children. This dataset is intended to facilitate research on visual attention modelling as it relates to ASD diagnosis. In the original study, the authors gathered eye movement data from a cohort comprising 14 children with ASD and 14 children with TD, all aged between 5 and 12 years, with a mean age of approximately 8 years. Notably, the dataset does not include subject identifiers distinguishing individuals within each group. Consequently, in alignment with previous work [4], fixation sequences indexed by i within the same group are assumed to originate from the same individual.

4.2 Experimental Setup

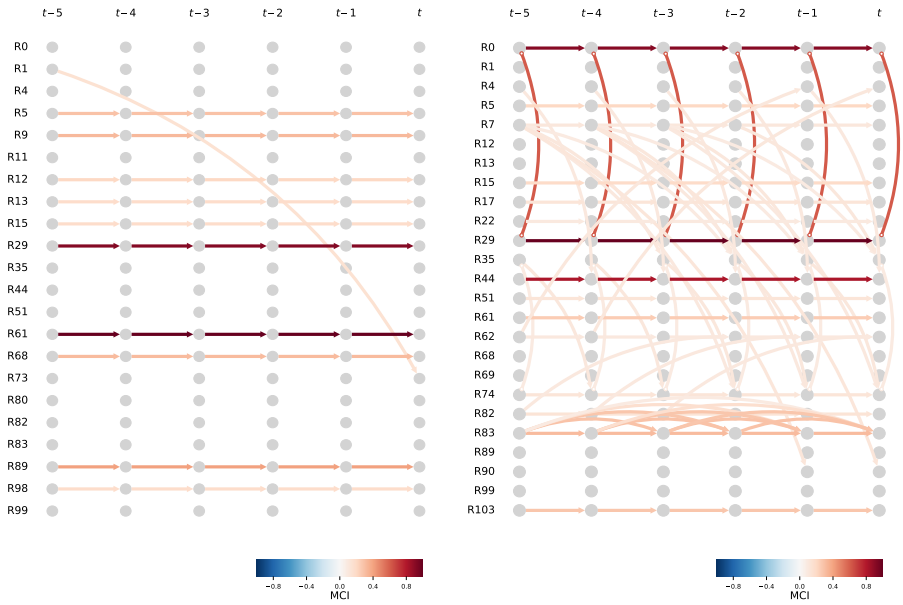
Each image from the Saliency4ASD dataset is automatically processed to extract relevant items within the scene. The SAM 2.1 Large model [23] is employed for

accurate segmentation of areas of interest, which are considered as regions of gaze attraction (items).

Aggregate item-level attention allocation traces (see Section 3.1) are generated by integrating the amount of attention (gaze points) that fall within the segmented regions. These traces are then fed into the Visual Attention Causal Discovery method outlined in Section 3. Items that do not attract attention are excluded from the computation of the Visual Attention Causal Graph.

Consistent with previous work [4, 32], the classification task is performed using a leave-one-subject-out strategy on 100 randomly selected images.

For the computation of the Visual Attention Causal Graph using the PCMCI method, we set the significance threshold to $\alpha = 0.01$ and define a time lag range of $\tau_{min} = 0$ and $\tau_{max} = 5$. Given the univariate and discrete/categorical nature of the variables in the Aggregated Item-Level Attention Allocation Traces, the G-squared Conditional Independence test is adopted [24].



(a) TD-group **(b)** ASD-group

Fig. 3: Time unrolled VA-CGs estimated on the same sample image from the TD (a) and ASD (b) groups providing a causal representation of visual attention dynamics across semantically segmented regions. Each node represents a region of interest (e.g., R0–R103) at successive time steps (t–5 to t). Directed edges indicate causal transitions of attention between regions, with edge colour denoting the causal strength as estimated by the Momentary Conditional Independence (MCI) test. The TD-group displays more consistent, linear patterns of attention allocation, while the ASD-group exhibits more complex and diverse transitions, indicating group-level differences in visual attention behaviour.

4.3 Results

We compare the proposed method against [4] and [32] and evaluate it in terms of accuracy, precision, recall, and AUC. Results are reported in Table 1.

Table 1: Comparison of the proposed method with state-of-the-art competing approaches on the Saliency4ASD dataset.

Methods	Accuracy	Precision	Recall	AUC
Chen et al. (Independent) [4]	0.89	0.86	0.93	0.92
Chen et al. (Full) [4]	**0.93**	0.93	0.93	0.98
AttentionGraph+S_{scan} [32]	**0.93**	-	-	-
AttentionGraph+S'_{scan} [32]	0.86	-	-	-
Our	**0.93**	0.93	0.93	0.93

Among these methods, our approach demonstrates strong and well-rounded performance, achieving an accuracy, precision, recall, and AUC of 0.93. This places it among the top-performing models, on par with Chen et al. (Full) [4] and AttentionGraph with S_{scan} in terms of accuracy. In contrast, Chen et al. (Independent) [4] shows slightly lower performance, with an accuracy of 0.89 and an AUC of 0.92, while AttentionGraph with S'_{scan} underperforms with an accuracy of 0.86.

It is worth noting that, unlike the purely predictive and, to some extent, black-box approach proposed in [4], the solution presented here offers a high degree of interpretability. This is achieved through the use of causal discovery techniques, which enable the method to characterise the underlying causal mechanisms governing the dynamics of attention allocation. By constructing Visual Attention Causal Graphs (VA-CGs), the proposed method not only predicts group membership accurately but also provides insights into how and why visual attention shifts over time within and across observer groups.

In this respect, the proposed solution shares some conceptual similarities with the AttentionGraph approach introduced in [32]; however, their method remains at a purely correlational level of explanation. In contrast, the adoption of VA-CGs allows the proposed approach to operate at a higher level on the ladder of causation [21], enabling the inspection of attention allocation mechanisms in neurodivergent subjects with greater explanatory depth.

As illustrated in Fig. 3, the time-unrolled VA-CGs offer an interpretable causal representation of visual attention dynamics across semantically segmented regions. The figure captures the temporal evolution of attention across regions of interest, with directed edges signifying causal transitions between them, and edge colour reflecting causal strength. Notably, the TD-group exhibits more consistent, linear attention patterns, whereas the ASD-group demonstrates more complex and varied transitions, highlighting group-level differences in visual attention behaviour. This makes the approach particularly valuable in domains

where understanding the decision-making process is as important as the predictions themselves, such as in cognitive and clinical research.

Overall, the proposed approach stands out for its balanced and complete performance combining high accuracy with strong precision, recall, and AUC, while offering substantial explanatory power.

5 Conclusion

In this work, we introduced a novel approach to model and analyse the causal mechanisms underlying visual attention allocation in individuals with and without Autism Spectrum Disorder (ASD). Our approach relies on the adoption of cutting-edge causal discovery methods for multivariate time-series data to construct directed graphs, referred to as Visual Attention Causal Graphs (VA-CGs), that capture how attention dynamically shifts between semantically relevant items in a visual scene. Experiments on the publicly available Saliency4ASD dataset demonstrate the potential of VA-CGs to accurately classify individuals as either ASD or typically developing (TD), achieving competitive performance compared to existing methods. Crucially, beyond mere classification, VA-CGs offer a significant advantage in terms of interpretability. By explicitly modeling causal relationships, the proposed framework provides insights into the distinct attentional strategies employed by different groups. This capability may be essential for advancing our understanding of the cognitive processes that differentiate ASD from TD individuals. Future work will explore the application of VA-CGs to other neurodevelopmental conditions (e.g., ADHD), as well as their potential for characterizing behavioral biomarkers in contexts such as personality trait prediction and biometric identification.

Acknowledgments

This work was supported by the PNRR project "Italian Strengthening of Esfri RI Resilience (ITSERR)" funded by the European Union - NextGenerationEU (CUP B53C22001770006).

References

1. Boccignone, G., Cuculo, V., D'Amelio, A., Grossi, G., Lanzarotti, R.: On gaze deployment to audio-visual cues of social interactions. IEEE Access **8**, 161630–161654 (2020)
2. Cartella, G., Cornia, M., Cuculo, V., D'Amelio, A., Zanca, D., Boccignone, G., Cucchiara, R.: Trends, Applications, and Challenges in Human Attention Modelling. In: IJCAI (2024)
3. Cartella, G., Cuculo, V., Cornia, M., Cucchiara, R.: Unveiling the Truth: Exploring Human Gaze Patterns in Fake Images. IEEE Signal Processing Letters **31**, 820–824 (2024)

4. Chen, S., Zhao, Q.: Attention-based autism spectrum disorder screening with privileged modality. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1181–1190 (2019)
5. Chen, X., Jiang, M., Zhao, Q.: Beyond average: Individualized visual scanpath prediction. In: CVPR (2024)
6. Corbetta, M., Shulman, G.L.: Control of goal-directed and stimulus-driven attention in the brain. Nature reviews neuroscience **3**(3), 201–215 (2002)
7. Cuculo, V., D'Amelio, A., Lanzarotti, R., Boccignone, G.: Personality gaze patterns unveiled via automatic relevance determination. In: Federation of International Conferences on Software Technologies: Applications and Foundations. pp. 171–184. Springer (2018)
8. Deng, S., Prasse, P., Reich, D.R., Dziemian, S., Stegenwallner-Schütz, M., Krakowczyk, D., Makowski, S., Langer, N., Scheffer, T., Jäger, L.A.: Detection of adhd based on eye movements during natural viewing. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. pp. 403–418. Springer (2022)
9. Desimone, R., Duncan, J.: Neural mechanisms of selective visual attention. Annual review of neuroscience **18**(1), 193–222 (1995)
10. D'Amelio, A., Patania, S., Bursic, S., Cuculo, V., Boccignone, G.: Using gaze for behavioural biometrics. Sensors **23**(3), 1262 (2023)
11. Friston, K., Adams, R.A., Perrinet, L., Breakspear, M.: Perceptions as hypotheses: saccades as experiments. Frontiers in psychology **3**, 151 (2012)
12. Frye, R.E., Vassall, S., Kaur, G., Lewis, C., Karim, M., Rossignol, D.: Emerging biomarkers in autism spectrum disorder: a systematic review. Annals of translational medicine **7**(23), 792 (2019)
13. Gerstenberg, T., Peterson, M.F., Goodman, N.D., Lagnado, D.A., Tenenbaum, J.B.: Eye-tracking causality. Psychological science **28**(12), 1731–1744 (2017)
14. Gutiérrez, J., Che, Z., Zhai, G., Le Callet, P.: Saliency4asd: Challenge, dataset and tools for visual attention modeling for autism spectrum disorder. Signal Processing: Image Communication **92**, 116092 (2021)
15. Ioannou, C., Seernani, D., Stefanou, M.E., Riedel, A., Tebartz van Elst, L., Smyrnis, N., Fleischhaker, C., Biscaldi-Schaefer, M., Boccignone, G., Klein, C.: Comorbidity matters: social visual attention in a comparative study of autism spectrum disorder, attention-deficit/hyperactivity disorder and their comorbidity. Frontiers in psychiatry **11**, 545567 (2020)
16. Itti, L., Koch, C.: Computational modelling of visual attention. Nature Reviews - Neuroscience **2**, 1–11 (2001)
17. Jiang, M., Zhao, Q.: Learning visual attention to identify people with autism spectrum disorder. In: Proceedings of the ieee international conference on computer vision. pp. 3267–3276 (2017)
18. Lan, G., Heit, B., Scargill, T., Gorlatova, M.: Gazegraph: Graph-based few-shot cognitive context sensing from human visual behavior. In: Proceedings of the 18th Conference on Embedded Networked Sensor Systems. pp. 422–435 (2020)
19. Liu, W., Li, M., Yi, L.: Identifying children with autism spectrum disorder based on their face processing abnormality: A machine learning framework. Autism Research **9**(8), 888–898 (2016)
20. Lumetti, L., Pipoli, V., Bolelli, F., Ficarra, E., Grana, C.: Enhancing patch-based learning for the segmentation of the mandibular canal. IEEE Access (2024)
21. Pearl, J.: Causality. Cambridge university press (2009)

22. Pelphrey, K.A., Sasson, N.J., Reznick, J.S., Paul, G., Goldman, B.D., Piven, J.: Visual scanning of faces in autism. Journal of autism and developmental disorders **32**, 249–261 (2002)
23. Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., et al.: Sam 2: Segment anything in images and videos. arXiv preprint arXiv:2408.00714 (2024)
24. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., Sejdinovic, D.: Detecting and quantifying causal associations in large nonlinear time series datasets. Science advances **5**(11), eaau4996 (2019)
25. Sasson, N.J., Turner-Brown, L.M., Holtzclaw, T.N., Lam, K.S., Bodfish, J.W.: Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. Autism Research **1**(1), 31–42 (2008)
26. Schiatti, L., Gori, M., Schrimpf, M., Cappagli, G., Morelli, F., Signorini, S., Katz, B., Barbu, A.: Modeling visual impairments with artificial neural networks: a review. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1987–1999 (2023)
27. Schütt, H.H., Rothkegel, L.O., Trukenbrod, H.A., Reich, S., Wichmann, F.A., Engbert, R.: Likelihood-based parameter estimation and comparison of dynamical cognitive models. Psychological review **124**(4), 505 (2017)
28. Spirtes, P., Glymour, C., Scheines, R.: Causation, prediction, and search. MIT press (2001)
29. Van Rijn, S., Urbanus, E., Swaab, H.: Eyetracking measures of social attention in young children: How gaze patterns translate to real-life social behaviors. Social Development **28**(3), 564–580 (2019)
30. Wang, S., Jiang, M., Duchesne, X.M., Laugeson, E.A., Kennedy, D.P., Adolphs, R., Zhao, Q.: Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. Neuron **88**(3), 604–616 (2015)
31. Wang, T., Xue, Y., Mohamed, Z.A., Jia, F.: Developmental functional brain network abnormalities in autism spectrum disorder comorbid with attention deficit hyperactivity disorder. European Journal of Pediatrics **184**(2), 166 (2025)
32. Yang, K.F., Li, Y.J.: Visual attention graph. arXiv preprint arXiv:2503.08531 (2025)
33. Zelinsky, G.J., Bisley, J.W.: The what, where, and why of priority maps and their interactions with visual working memory. Annals of the new York Academy of Sciences **1339**(1), 154–164 (2015)