Comparing eye-tracking time series imaging methods in ADHD screening

Pedro Lencastre[1,2], Ananta Devkota[1], **Shailendra Bhandari**[1,2], Bineeth Kuriakose[1], Pedro G. Lind[1,2,3,4]

[1]OsloMet – Oslo Metropolitan University; [2]OsloMet AI Lab, Oslo, Norway
[3]Kristiania University of Applied Sciences, Oslo, Norway
[4]Simula Research Laboratory, Oslo, Norway

30 June-5 July, 2025



INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS
IJCNN2025
30 JUNE - 5 JULY 2025 | ROME, ITALY
INTERNATIONAL NEURAL NETWORK SOCIETY

OSLOMET
OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

Introduction: ADHD and eye-tracking

- **Challenge**: ADHD diagnosis is complex, with up to 20% misdiagnosis rates [1].

[1] Elder, T. E. (2010)

Introduction: ADHD and eye-tracking

- **Challenge**: ADHD diagnosis is complex, with up to 20% misdiagnosis rates [1].
- **Eye-Tracking potential**: Non-invasive tool to capture gaze patterns (fixations, saccades) for ADHD screening.

[1] Elder, T. E. (2010)

Introduction: ADHD and eye-tracking

- **Challenge**: ADHD diagnosis is complex, with up to 20% misdiagnosis rates [1].
- **Eye-Tracking potential**: Non-invasive tool to capture gaze patterns (fixations, saccades) for ADHD screening.
- **Objective**: Compare time-series imaging methods (GASF, GADF, MTF, RP) using CNN for ADHD classification.

[1] Elder, T. E. (2010)

Dataset description

- **Source**: Healthy Brain Network (HBN) project, 1246 participants (568 ADHD, 152 TD), ages 6–21[2].

[2] O'Connor, D., et al., Gigascience, 6(2):giw011, 2017

Dataset description

- **Source**: Healthy Brain Network (HBN) project, 1246 participants (568 ADHD, 152 TD), ages 6–21[2].
- **Stimuli**: Four videos, "Fun with Fractals (2:44)", "Despicable Me (2:50)", "The Present (3:27)", and "Diary of a Wimpy Kid (1:57)", randomly ordered.

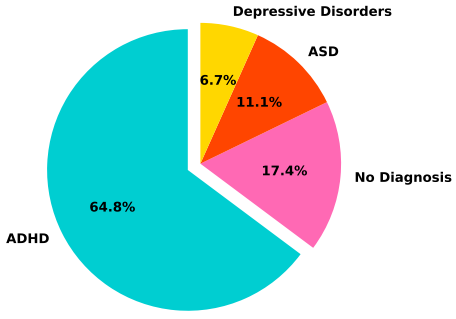
[2] O'Connor, D., et al., Gigascience, 6(2):giw011, 2017

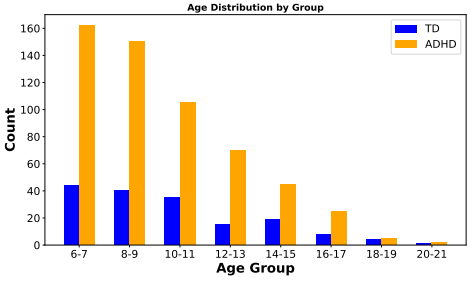
Dataset description

- **Source**: Healthy Brain Network (HBN) project, 1246 participants (568 ADHD, 152 TD), ages 6–21[2].
- **Stimuli**: Four videos, "Fun with Fractals (2:44)", "Despicable Me (2:50)", "The Present (3:27)", and "Diary of a Wimpy Kid (1:57)", randomly ordered.
- **Eye-Tracking**: Infrared tracker for right-eye gaze data (iView-X Red-m, SensoMotoric Instruments [SMI] GmbH, spatial resolution: 0.1°, accuracy: 0.5°, 60-120Hz).

[2] O'Connor, D., et al., Gigascience, 6(2):giw011, 2017

Dataset description

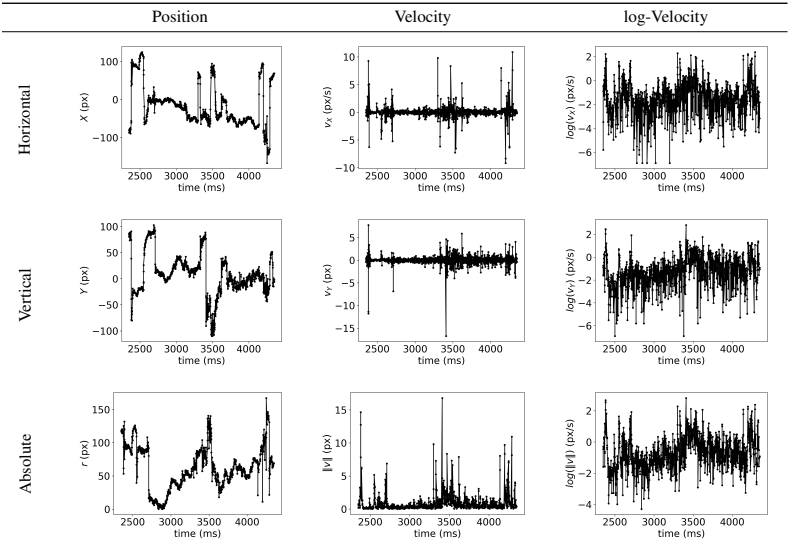




Eye-tracking time-series variables

- **Variables**: Gaze positions (X_t, Y_t, r_t), velocities ($v_x(t)$, $v_y(t)$, $v_s(t)$), and logarithmic velocities ($\log(|v_x(t)|)$, $\log(|v_y(t)|)$, $\log(|v_s(t)|)$).
- **Purpose**: Capture dynamic gaze patterns (fixations, saccades).

Eye-tracking time-series variables



Time-series imaging methods

- **Methods**: Gramian Angular Summation/Difference Fields (GASF/GADF), Markov Transition Field (MTF), Recurrence Plot (RP).

Time-series imaging methods

- **Methods**: Gramian Angular Summation/Difference Fields (GASF/GADF), Markov Transition Field (MTF), Recurrence Plot (RP).
- **Goal**: Transform time-series into images for CNN classification.

Time-series imaging methods

- **Methods**: Gramian Angular Summation/Difference Fields (GASF/GADF), Markov Transition Field (MTF), Recurrence Plot (RP).
- **Goal**: Transform time-series into images for CNN classification.
- 13176 images both ADHD and TD, with 80:10:10, 10674 train, 1185 validation, and 1317 test.

Gramian Angular Field (GAF) technique

- Introduces GASF and GADF to encode time series into images.

Gramian Angular Field (GAF) technique

- Introduces GASF and GADF to encode time series into images.
- Time series normalized to [-1, 1] using $\tilde{x}_i = \frac{x_i - \max(X) + (\min(X))}{\max(X) - \min(X)}$.

Gramian Angular Field (GAF) technique

- Introduces GASF and GADF to encode time series into images.
- Time series normalized to [-1, 1] using $\tilde{x}_i = \frac{x_i - \max(X) + (\min(X))}{\max(X) - \min(X)}$.
- Polar coordinates computed: $\phi_i = \cos^{-1}(\tilde{x}_i)$, $r_i = \frac{t_i}{N}$.

Gramian Angular Field (GAF) technique

- Introduces GASF and GADF to encode time series into images.
- Time series normalized to [-1, 1] using $\tilde{x}_i = \frac{x_i - \max(X) + (\min(X))}{\max(X) - \min(X)}$.
- Polar coordinates computed: $\phi_i = \cos^{-1}(\tilde{x}_i)$, $r_i = \frac{t_i}{N}$.
- GASF matrix: $\text{GASF}_{ij} = \cos(\phi_i + \phi_j)$, captures temporal correlations.

Gramian Angular Field (GAF) technique

- Introduces GASF and GADF to encode time series into images.
- Time series normalized to [-1, 1] using $\tilde{x}_i = \frac{x_i - \max(X) + (\min(X))}{\max(X) - \min(X)}$.
- Polar coordinates computed: $\phi_i = \cos^{-1}(\tilde{x}_i)$, $r_i = \frac{t_i}{N}$.
- GASF matrix: $\text{GASF}_{ij} = \cos(\phi_i + \phi_j)$, captures temporal correlations.
- GADF matrix: $\text{GADF}_{ij} = \sin(\phi_i - \phi_j)$, highlights differences.

Gramian Angular Field (GAF) technique

- Introduces GASF and GADF to encode time series into images.
- Time series normalized to [-1, 1] using $\tilde{x}_i = \frac{x_i - \max(X) + (\min(X))}{\max(X) - \min(X)}$.
- Polar coordinates computed: $\phi_i = \cos^{-1}(\tilde{x}_i)$, $r_i = \frac{t_i}{N}$.
- GASF matrix: $\text{GASF}_{ij} = \cos(\phi_i + \phi_j)$, captures temporal correlations.
- GADF matrix: $\text{GADF}_{ij} = \sin(\phi_i - \phi_j)$, highlights differences.

Matrix Representations

$$\text{GASF} = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \cdots & \cos(\phi_1 + \phi_n) \\ \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \cdots & \cos(\phi_n + \phi_n) \end{bmatrix}$$

$$\text{GADF} = \begin{bmatrix} \sin(\phi_1 - \phi_1) & \cdots & \sin(\phi_1 - \phi_n) \\ \vdots & \ddots & \vdots \\ \sin(\phi_n - \phi_1) & \cdots & \sin(\phi_n - \phi_n) \end{bmatrix}$$

Markov Transition Field (MTF)

- MTF transforms time series into images using Markov transition probabilities.

Markov Transition Field (MTF)

- MTF transforms time series into images using Markov transition probabilities.
- Process:
  1. Time series X_t follows Markov property:

$$\Pr(X_{t=j} = \hat{x}_j \mid X_{t=j-1} = \hat{x}_{j-1}, \ldots) = \Pr(X_{t=j} = \hat{x}_j \mid X_{t=j-1} = \hat{x}_{j-1})$$

  2. Discretize time series into N_s states, forming intervals (k_i, k_{i+1}).
  3. Construct transition matrix $\mathbf{T}$:

$$T_{i,j} = \Pr(X_{t=n+1} \in [k_i, k_{i+1}) \mid X_{t=n} \in [k_j, k_{j+1}))$$

  4. Represent $\mathbf{T}$ as an image (MTF).

Recurrence Plot (RP)

- RP visualizes recurring states in a time series as a 2D image.

Recurrence Plot (RP)

- RP visualizes recurring states in a time series as a 2D image.
- Maps time series to higher-dimensional phase space using sliding subsequences.

Recurrence Plot (RP)

- RP visualizes recurring states in a time series as a 2D image.
- Maps time series to higher-dimensional phase space using sliding subsequences.
- Process:
  1. Embed time series into phase space, where $x(i) \in \mathbb{R}^m$ is the i-th state.
  2. Compute distance matrix:

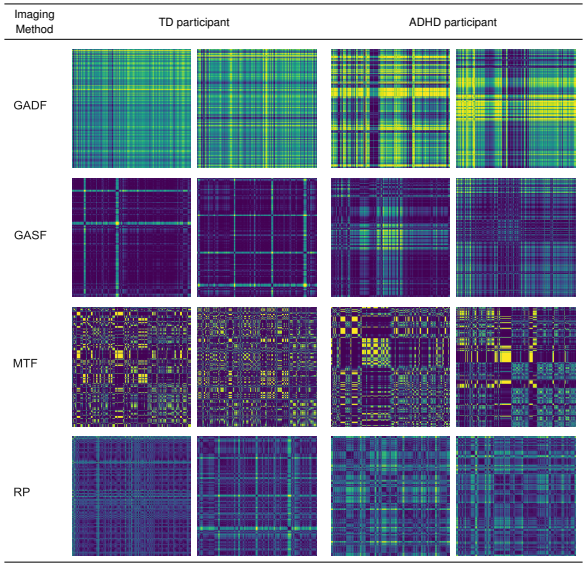
$$RP_{i,j} = \|x(i) - x(j)\|, \quad i, j = 1, \ldots, N$$

  3. Optionally binarize using threshold ε:

$$R_{i,j} = \Theta(\varepsilon - \|x(i) - x(j)\|)$$

  4. Plot distances as an image (RP).

Time-series imaging methods



Convolutional neural network (CNN)

- **Architecture**: Conv. layers (32, 64, 128, 256 filters), max-pooling, dense layers (512, 256, 128 units), dropout (0.5).

Convolutional neural network (CNN)

- **Architecture**: Conv. layers (32, 64, 128, 256 filters), max-pooling, dense layers (512, 256, 128 units), dropout (0.5).
- **Training**: 80% train, 10% validation, 10% test; 10-fold cross-validation.

Convolutional neural network (CNN)

- **Architecture**: Conv. layers (32, 64, 128, 256 filters), max-pooling, dense layers (512, 256, 128 units), dropout (0.5).
- **Training**: 80% train, 10% validation, 10% test; 10-fold cross-validation.

Table: Hyperparameter settings

Hyperparameters	**Value**
Optimizer	Adam
Loss function	Binary Cross-Entropy
Epochs	Up to 20 (Early stopping)
Batch size	32, 64, or 128
Learning rate	0.0001–0.01 (ReduceLROnPlateau)

Results: Classification accuracy

- **Key finding**: MTF achieves highest average accuracy (87.2%), excelling in 6/9 variables.
- **Others**: GASF best for $v_s(t)$ (92.3%); RP for $\log(|v_x(t)|)$ (86.2%).

Table: Classification accuracy across variables

Methods	Gaze position			Gaze velocity			Gaze velocity			Av. Acc.
	X_t	Y_t	r_t	$v_x(t)$	$v_y(t)$	$v_s(t)$	$\log(\|v_x(t)\|)$	$\log(\|v_y(t)\|)$	$\log(\|v_s(t)\|)$	
GADF	88.8	83.8	73.1	79.7	88.5	89.9	78.2	75.5	81.7	82.1
GASF	84.5	82.5	81.7	85.5	85.1	**92.3**	85.8	86.6	84.4	85.7
MTF	**89.9**	88.7	84.1	81.6	86.2	91.2	85.1	88.4	**89.7**	87.2
RP	89.6	84.8	78.0	81.1	82.2	89.7	**86.2**	83.7	85.6	84.5

Confusion matrices

- **Best performances**: MTF (X_t: 89.9%), GASF ($v_s(t)$: 92.3%), MTF ($\log(|v_s(t)|)$: 89.7%).
- **Observation**: Low false positives despite an imbalanced dataset (568 ADHD vs. 152 TD).

Table: Confusion matrices for best classifications

		Pred. label (MTF, X_t)		**Pred. label** (GASF, $v_s(t)$)		**Pred. label** (MTF, $\log(\|v_s(t)\|)$)	
		ADHD	TD	ADHD	TD	ADHD	TD
True label	ADHD	528	40	499	69	502	66
	TD	15	137	10	142	12	140

Summary and conclusions

- This study evaluates various time series imaging techniques for eye-tracking data.

Summary and conclusions

- This study evaluates various time series imaging techniques for eye-tracking data.
- Each technique excels with specific variables:
  - MTF: Best for horizontal eye movement (X) and best overall (87.2% avg. accuracy).
  - RP: Best for $\log(\|v_x(t)\|)$, (86.2%) accuracy.
  - GASF: Best for $v_s(t)$, (92.3%).
  - GADF: Best for $v_y(t)$.

Summary and conclusions

- This study evaluates various time series imaging techniques for eye-tracking data.
- Each technique excels with specific variables:
  - MTF: Best for horizontal eye movement (X) and best overall (87.2% avg. accuracy).
  - RP: Best for $\log(\|v_x(t)\|)$, (86.2%) accuracy.
  - GASF: Best for $v_s(t)$, (92.3%).
  - GADF: Best for $v_y(t)$.
- The imaging technique should be chosen based on the target variable.

Summary and conclusions

- **Implication**: Time-series imaging in eye-tracking is a promising, non-invasive method for ADHD screening.

Summary and conclusions

- **Implication**: Time-series imaging in eye-tracking is a promising, non-invasive method for ADHD screening.
- **Limitations**:
  - Dataset imbalance.
  - Only CNN classifier used.
  - No formal mathematical selection criteria for imaging methods.

Thank you

Thank you for your attention!



INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS
IJCNN2025
30 JUNE - 5 JULY 2025 | ROME, ITALY
INTERNATIONAL NEURAL NETWORK SOCIETY

OSLOMET
OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

simula

Funded by
The Research
Council of Norway