

Phân tích ảnh Eye Tracking để nhận dạng trẻ rối loạn phổ tự kỷ

Nguyễn Ngọc Hải, Tăng Tuyết Ngân, Nguyễn Thị Thủy Tiên, Trần Trung Tín*



Use your smartphone to scan this QR code and download this article

TÓM TẮT

Hiện nay, số lượng trẻ em mắc hội chứng rối loạn phổ tự kỷ (Autism Spectrum Diagnostic – ASD) ngày càng tăng trở thành vấn đề lớn của đời sống xã hội. Tự kỷ là một dạng rối loạn phát triển thần kinh gây nhiều khiếm khuyết về khả năng giao tiếp và tương tác xã hội của trẻ. Việc chẩn đoán lâm sàng trẻ mắc tự kỷ chủ yếu dựa trên chuyển động mắt của trẻ và kết hợp những bài kiểm tra về giáo dục và tâm lý. Chuyển động mắt đóng một vai trò quan trọng trong nhận thức và sự chú ý của cá nhân đối với các hoạt động xã hội. Các kỹ thuật phát hiện và theo dõi di chuyển của mắt không xâm lấn được phát triển trong nhiều thập kỷ gần đây. Ngày nay, phương pháp Eye Tracking cho phép đo lường và thu thập dữ liệu nhìn của mắt, nó được ứng dụng cho nhiều trong nghiên cứu về điều tra hành vi khác. Trong đó, ứng dụng phương pháp Eye Tracking nhận biết người mắc chứng rối loạn phổ tự kỷ được nhiều người quan tâm và nghiên cứu trong nhiều năm gần đây. Nghiên cứu này cung cấp cái nhìn tổng quát về sự chú ý thị giác của trẻ mắc rối loạn phổ tự kỷ (ASD) và trẻ phát triển bình thường (Typical Development - TD) với trọng tâm là tìm hiểu sự khác biệt trong hành vi quan sát của trẻ tự kỷ và trẻ bình thường. Nghiên cứu sử dụng thuật toán phân cụm DBSCAN (Density - Based Spatial Clustering of Applications with Noise) để phân vùng dữ liệu Eye Tracking được thu thập từ sự chú ý quan sát của trẻ trên các loại hình ảnh. Các thông số đặc trưng được rút ra từ bước phân tích dữ liệu là cơ sở để phân biệt sự khác nhau của hành vi quan sát hình ảnh giữa hai nhóm trẻ. Bên cạnh đó, mô hình mạng thần kinh nhân tạo Perceptron đa lớp (Multilayer Perceptron) trên cơ sở máy học chuyên sâu (Deep Learning) cũng được ứng dụng để phân loại và nhận dạng đặc trưng dữ liệu giữa hai nhóm trẻ. Bộ dữ liệu về chuyển động mắt của trẻ sử dụng cho nghiên cứu này được trích từ bài báo "A dataset of eye movements for the children with autism spectrum disorder" của tác giả Huiyu Duan và các cộng sự. Kết quả đạt được cho thấy triển vọng trong việc sử dụng dữ liệu Eye Tracking để nhận dạng trẻ tự kỷ từ hành vi quan sát hình ảnh của trẻ, đặc biệt là đối với các hình ảnh với chủ đề về con người.

Từ khoá: Rối loạn phổ tự kỷ, Eye Tracking, phân cụm DBSCAN, Deep Learning, Multilayer Perceptron

Khoa Khoa học Ứng dụng, Đại học Bách Khoa, Đại học Quốc gia Thành Phố Hồ Chí Minh, Việt Nam

Liên hệ

Trần Trung Tín, Khoa Khoa học Ứng dụng, Đại học Bách Khoa, Đại học Quốc gia Thành Phố Hồ Chí Minh, Việt Nam

Email: trtrtin@hcmut.edu.vn

Lịch sử

- Ngày nhận: 24-6-2021
- Ngày chấp nhận: 04-11-2021
- Ngày đăng: 25-12-2021

DOI : 10.32508/stdjet.v4i4.862



Bản quyền

© ĐHQG Tp.HCM. Đây là bài báo công bố mở được phát hành theo các điều khoản của the Creative Commons Attribution 4.0 International license.



TỔNG QUAN

Rối loạn phổ tự kỷ là tình trạng rối loạn phát triển ảnh hưởng nghiêm trọng đến giao tiếp ngôn ngữ, giao tiếp phi ngôn ngữ và tương tác xã hội kéo dài đến suốt đời¹. Mạng lưới theo dõi người khuyết tật của tổ chức CDC tại Mỹ, năm 2020 tỉ lệ là 1/54 trẻ được chẩn đoán mắc chứng rối loạn phổ tự kỷ (ASD)². Tỷ lệ trẻ em mắc tự kỷ gia tăng nhanh chóng đặt ra những vấn đề lớn đối với nhiều quốc gia. Ở Việt Nam giai đoạn 2000 – 2007, Nghiên cứu Mô hình tàn tật ở trẻ em của khoa Phục hồi Chức năng Bệnh viện Nhi Trung ương cho biết số lượng trẻ được chẩn đoán và điều trị tự kỷ ngày càng nhiều. Số lượng đến khám năm 2007 tăng gấp 50 lần so với năm 2000; xu thế mắc tự kỷ tăng nhanh từ 122% đến 268% trong giai đoạn 2004 – 2007 so với năm 2000³. Vào 1/2019, Tổng cục Thống kê công bố kết quả Điều tra Quốc gia về người khuyết tật tại Việt Nam hiện có khoảng 1 triệu người tự kỷ, tỉ lệ trẻ em mắc chứng tự kỷ ước tính là 1% số trẻ em sinh ra⁴.

Trẻ tự kỷ dành ít thời gian hơn để quan sát khuôn mặt và có xu hướng nhìn vào các mục tiêu không liên quan hoặc thường có xu hướng quan sát chi tiết, bỏ qua bức tranh tổng quát so với trẻ em phát triển bình thường (TD)⁵. Chính vì thế, một số thiết bị công nghệ đã ra đời để bắt được ánh nhìn của trẻ em nhằm mục đích sàng lọc sớm người mắc tự kỷ từ hai tháng tuổi trở lên. Hiện nay, thiết bị được đưa vào nghiên cứu và sử dụng tại một số phòng khám ở nước ngoài là mô hình Eye Tracking. Eye tracking là quy trình công nghệ ghi lại chuyển động của ánh mắt và thường được sử dụng nhất trong việc quan sát video hoặc ảnh có khuôn mặt người. Năm 2002, Pelphrey và các cộng sự đã nghiên cứu đầu tiên sử dụng Eye Tracking trong tự kỷ. Ông đã đưa ra kết luận rằng, các đối tượng mắc chứng tự kỷ đã dành ít thời gian hơn so với người bình thường để kiểm tra ở các đặc điểm cốt lõi của khuôn mặt (mắt, mũi và miệng)⁶. Năm 2002, Klin và các cộng sự cho rằng, những người mắc chứng tự kỷ có sự quan sát ở miệng nhiều hơn và họ dựa vào miệng

Trích dẫn bài báo này: Hải N N, Ngân T T, Tiên N T T, Tín T T. Phân tích ảnh Eye Tracking để nhận dạng trẻ rối loạn phổ tự kỷ. *Sci. Tech. Dev. J. - Eng. Tech.*; 4(4):1201-1211.

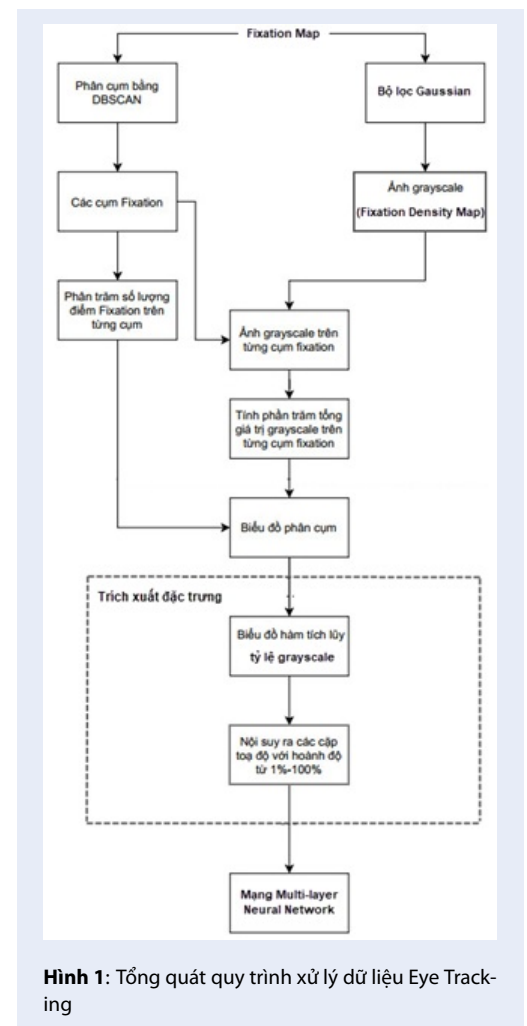
để xác định cảm xúc của người khác⁷. Năm 2006, Merin và các cộng sự kết luận, trẻ sơ sinh có sự suy giảm ánh nhìn ở mắt và nhìn chăm chăm vào miệng của các bà mẹ⁸. Năm 2007, Spezio và các cộng sự đã sử dụng Eye Tracking cùng với một phương pháp mới để thực hiện các kích thích, nhằm điều tra về khuôn mặt mà đối tượng đang quan sát để nhận ra biểu cảm cảm xúc⁹. Năm 2018, nghiên cứu Yanera và các cộng sự đã khẳng định người tự kỷ tập trung vào miệng của một người hay các đặc điểm khác hơn là mắt¹⁰. Năm 2016, Frazier và cộng sự đã tuyên bố rằng có sự khác biệt giữa hai nhóm trẻ em về thời gian nhìn tập trung cho kích thích xã hội và phi xã hội ở Khu vực quan tâm¹¹. Những nghiên cứu này, đã chứng minh Eye Tracking là thước đo tiêu chuẩn vàng về việc chẩn đoán tự kỷ thông qua đánh giá việc quan sát ánh nhìn. Trong nghiên cứu này, nhóm tác giả sử dụng bộ dữ liệu về chuyển động mắt của trẻ em được trích từ bài báo “A dataset of eye movements for the children with autism spectrum disorder” đã được tác giả Huiyu Duan (Harvard Medical School) và các cộng sự ở trường đại học Shanghai Jiao Tong University công bố trên MMSys '19: Proceedings of the 10th ACM Multimedia Systems Conference vào tháng 6/2019¹². Bên cạnh đó, sức mạnh của công nghệ trí tuệ nhân tạo (Artificial Intelligence) ngày càng phát triển nên phương pháp máy học chuyên sâu (Deep Learning) dần dần được đưa vào nghiên cứu nhận biết tự kỷ. Một số nghiên cứu¹³⁻¹⁶ chủ yếu tập trung vào cách chọn các đặc trưng với mong muốn rút ngắn thời gian chẩn đoán. Việc trích xuất các đặc trưng liên quan đến việc quan sát của trẻ tự kỷ và việc áp dụng các phương pháp thông qua mô hình Deep Learning cụ thể là thuật toán Multilayer perceptron sẽ giúp tăng tốc độ thời gian nhận biết, đánh giá và cải thiện độ chính xác khi nhận dạng và phân loại giữa hai nhóm trẻ.

PHƯƠNG PHÁP NGHIÊN CỨU

Bộ dữ liệu

Bộ dữ liệu về chuyển động mắt của trẻ em được trích từ Nghiên cứu “A dataset of eye movements for the children with autism spectrum disorder” đã được tác giả Huiyu Duan (Harvard Medical School) và các cộng sự ở trường đại học Shanghai Jiao Tong University công bố trên MMSys '19: Proceedings of the 10th ACM Multimedia Systems Conference vào tháng 6/2019¹². Để phân tích các đặc điểm và sự khác biệt của chuyển động mắt ở trẻ nhóm ASD và trẻ nhóm TD dưới các kích thích thị giác khác nhau, tác giả đã thu thập 300 hình ảnh theo 7 chủ đề đa dạng. Khi tiến hành thí nghiệm, tác giả đã sử dụng Tobii T120 Eye

Tracker để hiển thị hình ảnh và ghi lại chuyển động của mắt của những trẻ tham gia thí nghiệm. Thiết bị theo dõi mắt có màn hình 17 inch với độ phân giải 1280 x 1024 (rộng x cao). Tốc độ lấy mẫu của Eye Tracker là 120 Hz và phạm vi theo dõi khoảng 65 cm tính từ Eye Tracker. Tác giả đã tuyển chọn ra 14 trẻ mắc chứng rối loạn phổ tự kỷ (ASD) đã được chẩn đoán theo các tiêu chí DSM-V về chứng tự kỷ và 14 trẻ em phát triển bình thường (TD) để đối chứng. Quy trình thí nghiệm yêu cầu những đứa trẻ nhìn tự do vào hình ảnh. Nhóm nghiên cứu đã sử dụng hết 300 hình ảnh số và dữ liệu Eye Tracking của 28 trẻ khi quan sát hình ảnh. Dữ liệu mà tác giả cung cấp bao gồm: tọa độ (x, y) và thời lượng quan sát tại mỗi điểm Fixation. Quy trình xử lý dữ liệu được thể hiện ở sơ đồ khối tại Hình 1.



Hình 1: Tổng quát quy trình xử lý dữ liệu Eye Tracking

Phương pháp phân cụm DBSCAN

Từ dữ liệu ta được các điểm Fixation của trẻ ở cả hai nhóm (TD và ASD) được biểu diễn dưới dạng một bản đồ nhị phân, trong đó giá trị 1 ở những vị trí có ghi nhận các điểm Fixation, giá trị 0 ở những vị trí không ghi nhận các điểm Fixation. Tiếp theo, bản đồ Fixation được làm mịn với bộ lọc Gaussian để tạo ra bản đồ mật độ Fixation (còn gọi là bản đồ chú ý trực quan). Ngoài ra, trong nghiên cứu này, bản đồ Fixation đã được bổ sung thêm dữ liệu ghi nhận thời lượng quan sát của trẻ tương ứng với tọa độ điểm Fixation, khác với FDM mà tác giả cung cấp. Điều này cho phép ảnh grayscale trên FDM càng thể hiện rõ ràng hơn mức độ chú ý của trẻ được tham gia khảo sát, những điểm Fixation có thời lượng nhìn càng lâu thì điểm Fixation đó và vùng xám được tạo bởi điểm Fixation đó trên FDM càng sáng.

Trên bản đồ Fixation Map, các điểm Fixation được phân cụm dựa trên mật độ của chúng trên ảnh bằng thuật toán phân cụm DBSCAN (Density-Based Spatial Clustering of Applications with Noise). Từ đây, ảnh sẽ được phân chia thành nhiều vùng dựa theo các vị trí được xác định bởi các cụm. Kết hợp với ảnh grayscale trên FDM, ta có những vùng grayscale cũng được xác định bởi các cụm. Tính tổng các giá trị grayscale trên mỗi vùng, ta được biểu đồ giá trị grayscale trên các vùng được xác định bởi các cụm. Khi đó, vùng được chú ý nhiều hơn thì cũng sẽ có giá trị grayscale lớn hơn.

- eps (ϵ): khoảng cách chỉ định các vùng lân cận, các điểm Fixation được xem là ở cùng một cụm nếu khoảng cách giữa chúng có giá trị nhỏ hơn hoặc bằng giá trị khoảng cách eps.
- MinPts: số lượng các phần tử ít nhất để tạo thành một cụm độc lập (Cluster). Trong nghiên cứu này, MinPts được quy ước có giá trị là 2, các cụm có số lượng phần tử ít hơn 2 được gọi là các cụm nhiễu (Noise).

Về việc xác định giá trị eps, ta gọi là bán kính của vùng tròn xám được tạo bởi điểm Fixation qua bước làm mịn bằng bộ lọc Gaussian. Tuy nhiên, sự phân bố của các điểm Fixation trên hình không đồng đều, các điểm nhiều có tính phân tán và riêng lẻ, trong khi các cụm có mật độ càng cao thì các điểm Fixation có khuynh hướng chồng chập lẫn nhau. Khi đó, khoảng cách chỉ định các vùng lân cận cũng sẽ có giá trị khác nhau. Do vậy, nhóm nghiên cứu đã đề xuất khái niệm khoảng cách chỉ định trung bình được xác định theo công thức:

$$\bar{\epsilon} = \sqrt{\frac{A}{\pi N}}$$

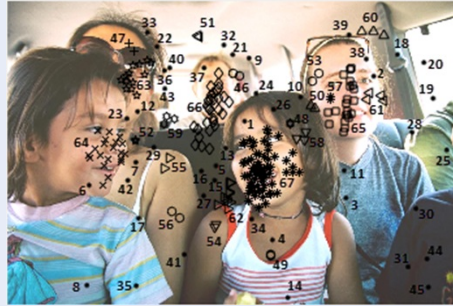
Với: A là tổng diện tích các vùng tròn xám được tạo bởi các điểm Fixation qua bước làm mịn bằng bộ lọc Gaussian và N là tổng số điểm Fixation trên mỗi ảnh. Từ bộ dữ liệu của hai nhóm trẻ (TD và ASD), ta thu được hai tập hợp điểm Fixation của từng nhóm trẻ được minh họa ở Hình 2. Màu sắc của điểm Fixation của hai nhóm có màu càng đậm khi thời lượng nhìn của trẻ càng lâu.



Tiến hành phân cụm bằng phương pháp DBSCAN cho toàn bộ các điểm Fixation trên ảnh của cả hai nhóm trẻ (TD và ASD), kết quả thu được là các cụm (Cluster) có mật độ phân bố các điểm Fixation khác nhau, các cụm được thể hiện bằng ký tự đặc trưng, có gắn nhãn số thứ tự cụm và có thể chứa cùng lúc điểm Fixation của cả hai nhóm trẻ. Thứ tự các cụm được quy ước sắp xếp tăng dần theo số lượng điểm Fixation thuộc cụm. Như vậy, các điểm nhiều sẽ được xem là các cụm một phần tử và có số thứ tự thấp, các cụm nhiều phần tử thì sẽ là các vùng AOI (Area Of Interest) có số thứ tự cao (Hình 3).

Từ kết quả phân cụm DBSCAN, ta xác định vùng ảnh bao quanh cụm cho mỗi trường hợp. Số lượng trường hợp phân tích vùng ảnh bao quanh cụm bằng đúng số lượng các cụm Hình 4).

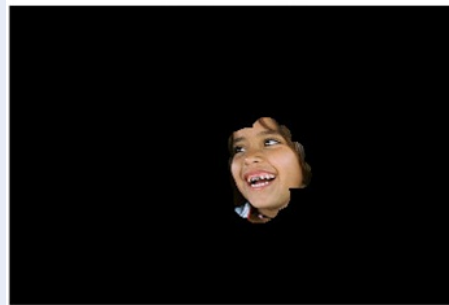
Tương tự, ta cũng xác định được vùng ảnh grayscale bao quanh cụm trên ảnh FDM của hai nhóm trẻ riêng



Hình 3: Kết quả phân cụm các điểm Fixation của hai nhóm trẻ bằng phương pháp DBSCAN



(a)

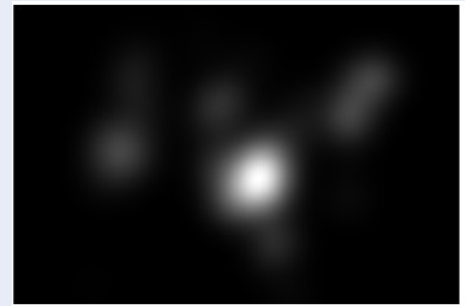


(b)

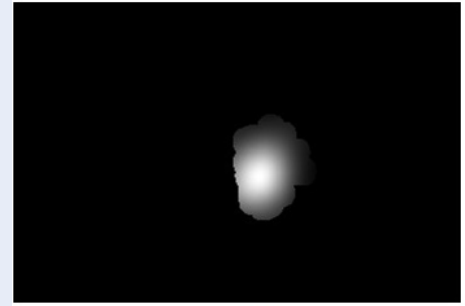
Hình 4: a) Các điểm Fixation thuộc cùng một cụm; b) Vùng ảnh bao quanh một cụm điểm Fixation

biệt (Hình 5 và Hình 6).

Ta có N vùng ảnh grayscale sau quá trình kết hợp ảnh FDM của mỗi nhóm trẻ với kết quả phân cụm DBSCAN. Vùng ảnh grayscale bao quanh cụm có diện tích càng rộng khi cụm có kích thước càng lớn. Bên cạnh đó, giá trị grayscale của vùng càng lớn khi trẻ nhìn vào vùng đó càng lâu. Trên mỗi vùng ảnh grayscale, ta tính tổng các giá trị grayscale của mỗi pixel (SOG – Sum Of Grayscale). Như vậy, ta có N giá trị SOG cho mỗi nhóm trẻ. Từ đó, ta thiết lập biểu đồ tỷ lệ grayscale của vùng bao quanh cụm (Chart Of the



(a)



(b)

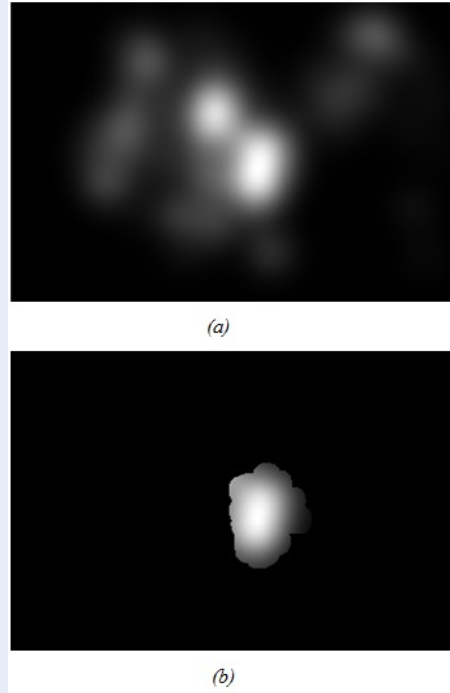
Hình 5: Ảnh FDM của trẻ phát triển bình thường (TD). a) Toàn bộ FDM; b) Vùng grayscale của FDM bao quanh một cụm

Grayscale Ratio on the Area Surrounding Cluster) với trục hoành là thứ tự các cụm, trục tung là tỷ lệ phần trăm grayscale của vùng bao quanh cụm (Hình 7).

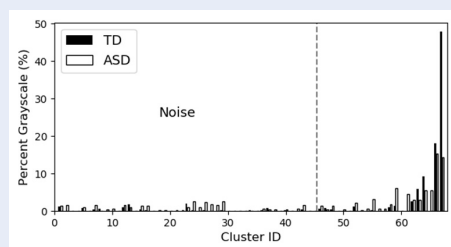
Trích xuất đặc trưng

Từ biểu đồ tỷ lệ grayscale trên vùng bao quanh cụm với quy luật tăng giảm khá phức tạp giữa các vùng, biểu đồ hàm tích lũy tỷ lệ grayscale với đặc điểm hàm là luôn luôn tăng được sử dụng như một công cụ đo lường tốc độ tăng trưởng giá trị tích lũy grayscale (Hình 8).

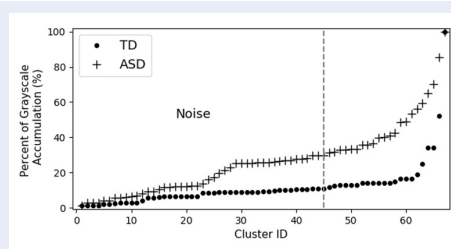
Tuy nhiên, do số lượng cụm giữa các hình ảnh là khác nhau. Để thuận tiện cho việc so sánh kết quả tăng trưởng hàm tích lũy tỷ lệ grayscale giữa các hình, thứ tự các cụm trên trục hoành của biểu đồ sẽ được thay thế bằng giá trị tích lũy tỷ lệ phần trăm số lượng các điểm Fixation của các cụm. Khi đó, biểu đồ tích lũy tỷ lệ grayscale của các ảnh từ bộ dữ liệu đều có trục hoành như nhau và luôn biến thiên từ 1% đến 100%. Kết quả thu được như Hình 9, qua quá trình khảo sát định tính, nhóm nghiên cứu nhận thấy rằng ở đa số các trường hợp khảo sát, trẻ phát triển bình thường hay nhìn tập trung vào khu vực trung tâm và có khuynh hướng hứng thú với các chi tiết nổi bật trên ảnh (vùng AOI). Trong khi đó, trẻ tự kỷ lại hay



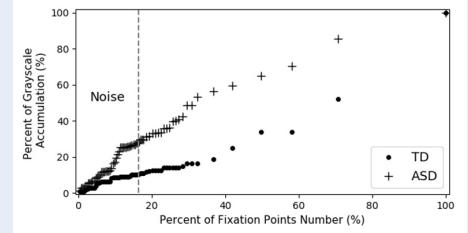
Hình 6: Ảnh FDM của trẻ rối loạn phổ tự kỷ. (ASD) a) Toàn bộ FDM; b) Vùng grayscale của FDM bao quanh một cụm



Hình 7: Biểu đồ tỷ lệ grayscale của vùng ảnh bao quanh cụm

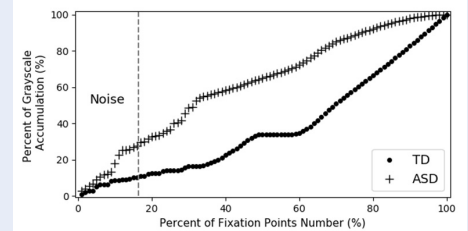


Hình 8: Hàm tích lũy tỷ lệ grayscale theo thứ tự các cụm



Hình 9: Hàm tích lũy tỷ lệ grayscale theo tích lũy phần trăm số lượng điểm Fixation

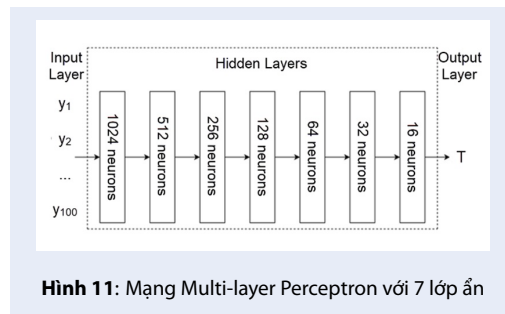
thiếu tập trung khi quan sát, có khuynh hướng nhìn vào các chi tiết nhỏ trên ảnh. Do đó, biểu đồ tích lũy tỷ lệ grayscale của trẻ tự kỷ có khuynh hướng tăng sớm ở các cụm thuộc miền nhiễu và chậm dần ở các vùng AOI phân bố ở miền bên phải của trục hoành. Ngược lại, trẻ phát triển bình thường có biểu đồ tích lũy grayscale tăng chậm ở miền nhiễu và tăng nhanh dần ở các vùng AOI. Có thể khẳng định rằng, biến thiên tăng trưởng của biểu đồ tích lũy tỷ lệ grayscale chính là dấu hiệu đặc trưng để nhận biết và phân biệt giữa hai nhóm trẻ. Tuy nhiên, khi áp dụng mô hình nhận dạng Neural Network thì đòi hỏi số lượng dữ liệu đầu vào (inputs) từ kết quả hàm tích lũy tỷ lệ grayscale giữa các ảnh là như nhau. Nhưng thực tế thì số lượng inputs giữa các ảnh là khác nhau. Do đó, kết quả hàm tích lũy tỷ lệ grayscale sẽ được áp dụng thêm phương pháp nội suy cubic để tạo số lượng dữ liệu đặc trưng như nhau cho tất cả các ảnh khảo sát. Trong nghiên cứu này, quá trình nội suy cubic được quy ước tạo ra 100 dữ liệu đặc trưng cho cả hai nhóm trên mỗi ảnh. Khi đó, dữ liệu đầu vào cho mô hình nhận dạng Neural Network luôn có số lượng cố định là 100 inputs, đó là tỷ lệ phần trăm tích lũy grayscale tương ứng với tỷ lệ phần trăm tích lũy số lượng điểm Fixation từ 1% đến 100%, theo quy tắc sắp xếp từ trái sang phải trên trục hoành là các cụm có số lượng điểm Fixation từ ít đến nhiều dần (Hình 10).



Hình 10: Biểu đồ hàm tích lũy tỷ lệ grayscale qua phép nội suy cubic

Phương pháp phân loại

Perceptron là một thuật toán phân loại cho trường hợp đơn giản nhất: chỉ có hai class (hay còn gọi là binary classification) và hoạt động được trong một số trường hợp rất cụ thể. Mạng Perceptron là nền tảng cho một mảng lớn quan trọng của Machine Learning là Neural Networks và sau này là Deep Learning¹⁷. Trong nghiên cứu này, mô hình Multilayer Perceptron được áp dụng để phân loại 100 đặc trưng từ kết quả của hàm tích lũy tỷ lệ grayscale cho cả trẻ bình thường và trẻ tự kỷ. Qua quá trình thử nghiệm, nhóm nghiên cứu đã sử dụng mô hình mạng Multi-layer Perceptron có 7 lớp ẩn với số lượng neuron theo sơ đồ Hình 11. Lớp ra có một neuron với $T = 0$ là nhân tương ứng với dữ liệu của trẻ bình thường hoặc $T = 1$ là nhân tương ứng với trẻ rối loạn tự kỷ. Hàm hoạt tính được sử dụng là hàm ReLU, được sử dụng khá nhiều trong những năm gần đây khi huấn luyện các mạng neuron với tốc độ tính toán và xử lý nhanh hơn các hàm sigmoid.



Hình 11: Mạng Multi-layer Perceptron với 7 lớp ẩn

KẾT QUẢ VÀ THẢO LUẬN

Sau khi tiến hành phân cụm bằng phương pháp DBSCAN và lập biểu đồ phân cụm. Cụ thể, hầu hết trẻ phát triển bình thường hay có khuynh hướng quan sát vào khu vực giữa của ảnh, trong khi sự quan sát của nhóm trẻ tự kỷ thì lại mang tính phân tán rộng trên ảnh, thể hiện sự thiếu tập trung. Bên cạnh đó, trẻ tự kỷ cũng dành nhiều thời gian hơn để quan sát những chi tiết hình nền (ví dụ như mặt biển, sân cỏ, ...) thay vì dành phần lớn thời gian để quan sát các chi tiết chính nổi bật trong ảnh như đối với trẻ bình thường. Điều này có thể lý giải rằng đa số trẻ tự kỷ có tính cách thích khám phá, tìm kiếm các đối tượng nhỏ nhất, ít nổi bật ở xung quanh. Qua quá trình phân tích kết quả đạt được, nghiên cứu rút ra được một vài nhận định như sau:

Đối với những hình ảnh có người và đồ vật, trẻ tự kỷ tuy vẫn chú ý đến khuôn mặt người nhưng ít hơn so với trẻ bình thường, và thích quan sát những đồ vật

khác hoặc các chi tiết nền xung quanh người. Biểu đồ tích lũy tỷ lệ grayscale ở miền nhiều của trẻ tự kỷ luôn tăng nhanh hơn nhiều so với trẻ bình thường đối với loại hình ảnh này. Trong khi đó, biểu đồ tích lũy tỷ lệ grayscale của trẻ bình thường thì lại tăng nhanh ở các vùng AOI được phân bố ở miền bên phải trên trục hoành (Hình 12).

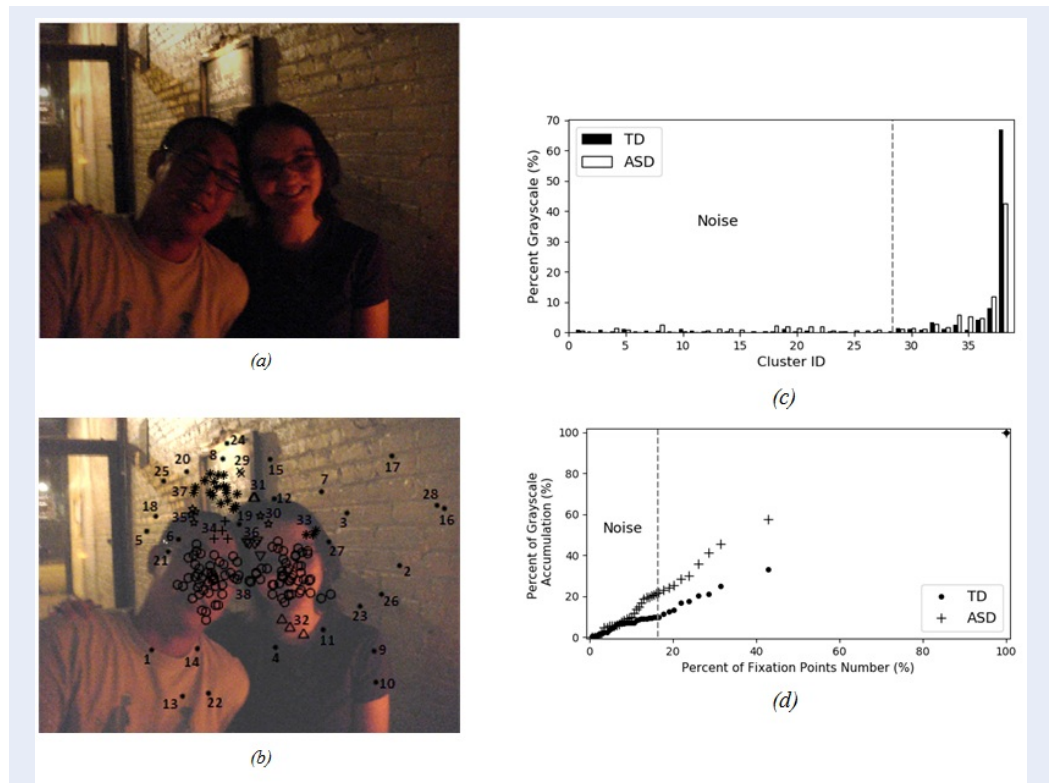
- Đối với ảnh động vật (Hình 13), trẻ tự kỷ có thể thoải mái quan sát hình ảnh khá giống như trẻ bình thường, các đối tượng quan sát như khuôn mặt, cơ thể động vật và các chi tiết xung quanh cũng không có sự khác nhau nhiều giữa hai nhóm trẻ.

Đối với hình thiên nhiên, dạng hình ảnh chỉ có sự vật mà không có người cũng tạo sự khác biệt giữa hai nhóm trẻ. Trẻ bình thường sẽ có khuynh hướng quan sát tập trung vào khu vực trung tâm hoặc sự vật nổi bật, đường đặc trưng của nhóm TD tăng nhanh ở các vùng AOI, đặc biệt là ở vùng AOI_{max}. Còn đối với nhóm trẻ tự kỷ, sự quan sát vẫn có tính tập trung vào khu vực trung tâm hoặc đồ vật nổi bật nhưng với thời lượng quan sát ít hơn. Với thời lượng còn lại, trẻ tự kỷ thường nhìn phân tán xung quanh với thói quen là thích tìm kiếm, khám phá các chi tiết nhỏ nhất. Do vậy, đường đặc trưng của nhóm ASD lại tăng khá nhanh so với nhóm TD ở các cụm nhiều (Hình 14). Khảo sát tập dữ liệu ảnh theo hai trường hợp:

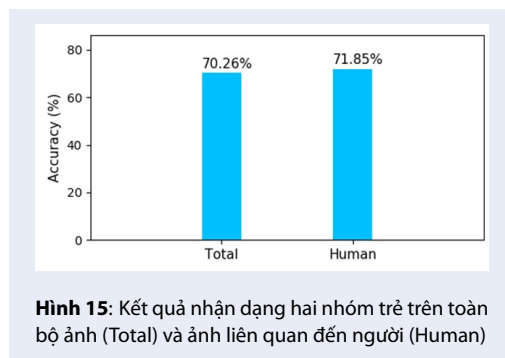
- Total: bao gồm toàn bộ 300 hình ảnh của tập dữ liệu với đầy đủ các chủ đề động vật, thiên nhiên, con người, sự vật, sự việc, hiện tượng, ...
- Human: 153 hình ảnh chỉ có người hoặc con người là đối tượng chính trong hình với các biểu hiện như gương mặt với nhiều cảm xúc, hành động hoặc vừa hành động vừa có biểu cảm khuôn mặt, ...

Sử dụng mạng Multilayer Perceptron cho hai trường hợp trên, kết quả nhận dạng thu được lần lượt là 71,85% và 70,26% như Hình 15. Tuy sự chênh lệch không nhiều nhưng có thể thấy được sự tác động của chủ đề về người mang lại ý nghĩa khác biệt đối với dữ liệu nhìn của trẻ bình thường và trẻ tự kỷ.

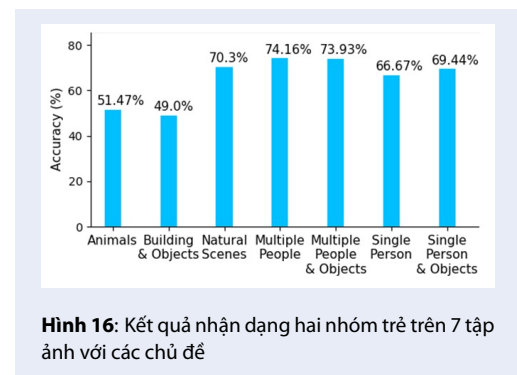
Từ tập dữ liệu 300 ảnh, việc phân loại chủ đề ảnh được tiến hành chi tiết hơn. Cụ thể trong nghiên cứu này đã phân ra 7 loại ảnh theo các chủ đề: Động vật (Animals), Tòa nhà và đồ vật (Buildings & Objects), Phong cảnh thiên nhiên (Natural Scenes), Nhiều người (Multiple People), Nhiều người và đồ vật (Multiple People & Objects), Một người (Single Person), Một người và đồ vật (Single Person & Objects). Kết quả nhận dạng, phân loại giữa hai nhóm trẻ qua



Hình 12: Kết quả phân tích đối với hình ảnh người và đồ vật. a) Ảnh gốc b) Kết quả phân cụm DBSCAN các điểm fixation trên ảnh c) Biểu đồ tỷ lệ grayscale d) Biểu đồ đặc trưng tích lũy tỷ lệ grayscale



Hình 15: Kết quả nhận dạng hai nhóm trẻ trên toàn bộ ảnh (Total) và ảnh liên quan đến người (Human)

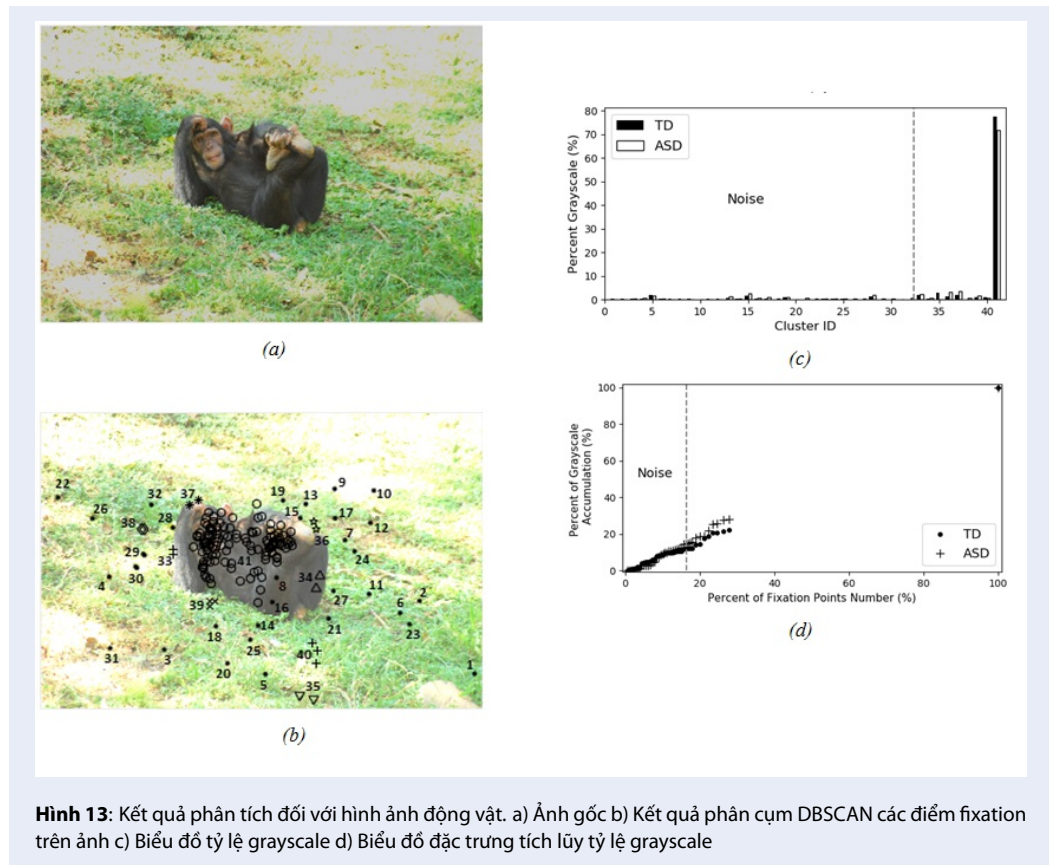


Hình 16: Kết quả nhận dạng hai nhóm trẻ trên 7 tập ảnh với các chủ đề

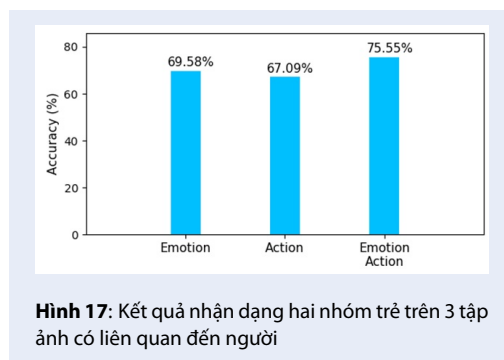
mạng Multi-layer Perceptron được trình bày ở biểu đồ cột Hình 16.

Phân tích kết quả đạt được, loại ảnh có tỷ lệ nhận dạng thấp là ảnh các đồ vật, tòa nhà và động vật; cao nhất đối với loại ảnh nhiều người. Kết quả nhận dạng thấp đối với loại ảnh động vật ta có thể dự đoán trước vì từ biểu đồ đặc trưng tích lũy tỷ lệ grayscale cho thấy không có sự khác biệt đáng kể về cách thức quan sát hình ảnh của hai nhóm trẻ. Bên cạnh đó, chủ đề về thiên nhiên đặc biệt cho kết quả nhận dạng khá cao trong nhóm loại ảnh không có con người, cho thấy

tiềm năng của loại ảnh này khi sử dụng nó làm kích thích trong các thí nghiệm Eye Tracking. Các điểm Fixation của nhóm trẻ tự kỷ rất phân tán, lý do vì trẻ tự kỷ gặp khó khăn trong việc tập trung¹⁸. Ngoài ra còn một nguyên nhân nữa là trẻ tự kỷ thường có cái nhìn cục bộ, có xu hướng quan sát chi tiết và bỏ qua bức tranh tổng thể⁵; điển hình như: “trẻ tự kỷ thấy cái cây trước khi thấy khu rừng”¹⁹. Trong khi đó, các loại ảnh liên quan đến con người thì đều đạt độ chính xác cao và xấp xỉ gần nhau. Vì vậy, các loại ảnh có



chứa yếu tố con người là một chủ đề có thể khai thác được rất nhiều từ các khía cạnh khác nhau. Cụ thể, từ nhóm loại ảnh về người lại được phân loại ra ba chủ đề: Emotion (Ảnh chỉ thể hiện cảm xúc), Action (Ảnh chỉ thể hiện hành động), Emotion & Action (Ảnh vừa thể hiện cảm xúc lẫn hành động). Kết quả nhận dạng hai nhóm trẻ qua mạng Multilayers Perceptron được trình bày ở Hình 17.

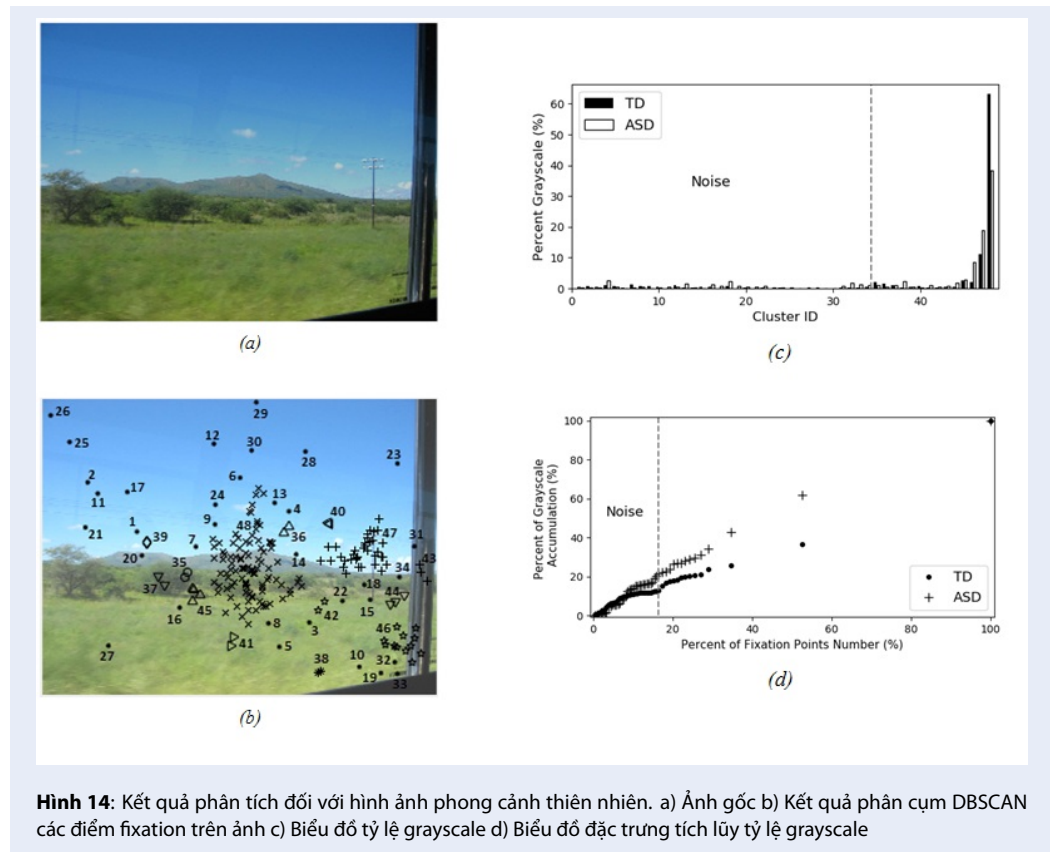


Phân tích kết quả đạt được cho thấy những bức ảnh vừa có cảm xúc vừa có hành động mang lại kết quả phân loại tốt nhất. Trẻ tự kỷ không thích tiếp xúc với

con người cũng như xã hội nên sẽ dành thời gian để quan sát vào các đồ vật xung quanh. Trong một số trường hợp, trẻ tự kỷ thích nhìn tay hoặc đồ vật được người đó cầm. Ngoài ra, chủ đề cảm xúc cũng cho kết quả phân loại tốt, kết quả này sẽ cao hơn nếu kích thích được sử dụng là ảnh động hoặc video.

KẾT LUẬN

Việc nhận dạng trẻ mắc rối loạn phổ tự kỷ bằng phương pháp phân tích ảnh Eye Tracking cho thấy có nhiều triển vọng. Từ các kết quả đạt được cho thấy việc áp dụng phương pháp phân cụm DBSCAN rất phù hợp với đặc tính phân bố của các điểm Fixation trên ảnh của cả hai nhóm trẻ. Bên cạnh đó, quá trình trích xuất đặc trưng của hàm tích lũy tỷ lệ grayscale cũng cho thấy được tiềm năng cho việc nhận dạng trẻ tự kỷ thông qua thuật toán Deep Learning. Nghiên cứu cũng đã trình bày ra được mức độ phù hợp của các loại ảnh khi sử dụng, nhằm tạo sự thuận lợi cho việc lựa chọn tối ưu chủ đề hình ảnh trong các nghiên cứu về sau liên quan đến việc nhận biết và chẩn đoán trẻ tự kỷ qua dữ liệu Eye Tracking. Tuy nhiên, cần nhiều nghiên cứu sâu hơn cũng như quy mô lớn hơn về tập dữ liệu Eye Tracking để phục vụ cho việc nhận



Hình 14: Kết quả phân tích đối với hình ảnh phong cảnh thiên nhiên. a) Ảnh gốc b) Kết quả phân cụm DBSCAN các điểm fixation trên ảnh c) Biểu đồ tỷ lệ grayscale d) Biểu đồ đặc trưng tích lũy tỷ lệ grayscale

dạng chính xác và chẩn đoán mức độ tự kỷ ở trẻ với chi phí chấp nhận được và quy trình thí nghiệm đơn giản, có khuôn mẫu và hệ thống.

LỜI CẢM ƠN

Nghiên cứu này được tài trợ bởi Trường Đại học Bách Khoa - ĐHQG TP.HCM trong khuôn khổ **đề tài mã số T-KHUD-2020-71**.

DANH MỤC CÁC TỪ VIẾT TẮT

ASD: Rối loạn phổ Tự kỷ – Autism Spectrum Disorder

TD: Phát triển bình thường – Typical Development

FDM: Bản đồ mật độ Fixation – Fixation Density Map

DBSCAN: Thuật toán phân cụm dựa trên mật độ – Density-Based Spatial Clustering of Applications with Noise

AOI: Vùng quan sát nổi bật (có nhiều điểm Fixation) – Area Of Interest

AOI_{max} : Vùng quan sát có số lượng điểm Fixation nhiều nhất.

SOG: Tổng các giá trị grayscale – Sum Of Grayscale

XUNG ĐỘT LỢI ÍCH

Nhóm tác giả cam kết rằng không có xung đột lợi ích khi thực hiện nghiên cứu này.

ĐÓNG GÓP CỦA CÁC TÁC GIẢ

Tất cả các tác giả đóng góp vào việc nghiên cứu và xử lý dữ liệu cũng đã đánh giá kết quả. Tất cả các tác giả đều tham gia giải thích dữ liệu, biên tập và hoàn thiện bản thảo.

TÀI LIỆU THAM KHẢO

- Đạt TQ, Quốc NTTT. Giảm biểu hiện gen ABCC4 (ATP-binding cassette subfamily C) liên quan đến rối loạn tự kỷ trên mô hình ruồi giấm. *Bản B của Tạp chí Khoa học và Công nghệ Việt Nam*. 2020;62(7);Available from: https://b.vjst.vn/index.php/ban_b/article/view/846.
- Speaks A. Autism statistics and facts. Retrieved September. 2020;16:2020;Available from: <https://www.autismspeaks.org/autism-statistics-asd>.
- Giang NTH, Hà TTT. Nghiên cứu xu thế mắc và một số đặc điểm dịch tễ học của trẻ tự kỷ điều trị tại bệnh viện Nhi Trung ương giai đoạn 2000 đến 2007. *Tạp chí Y học thực hành*. 2008;4:104-107;.
- Hiền T. Trẻ em tự kỷ tại Việt Nam gia tăng, phần lớn không được trị liệu, giáo dục phù hợp Hồ Chí Minh: Tuổi trẻ Online; 2021;Available from: <https://tuoitre.vn/tre-em-tu-ky-tai-viet-nam-gia-tang-phan-lon-khong-duoc-tri-lieu-giao-duc-phu-hop>.
- Happé F, Frith U. The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*. 2006;36(1):5-25;Available from: <https://doi.org/10.1007/s10803-005-0039-0>.

6. Pelphrey KA, Sasson NJ, Reznick JS, Paul G, Goldman BD, Piven J. Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*. 2002;32(4):249-261; Available from: <https://doi.org/10.1023/a:1016374617369>.
7. Klin A, Jones W, Schultz R, Volkmar F, Cohen D. Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of general psychiatry*. 2002;59(9):809-816; Available from: <https://doi.org/10.1001/archpsyc.59.9.809>.
8. Merin N, Young GS, Ozonoff S, Rogers SJ. Visual fixation patterns during reciprocal social interaction distinguish a subgroup of 6-month-old infants at-risk for autism from comparison infants. *Journal of Autism and Developmental Disorders*. 2007;37(1):108-121; Available from: <https://doi.org/10.1007/s10803-006-0342-4>.
9. Spezio ML, Adolphs R, Hurlley RS, Piven J. Abnormal use of facial information in high-functioning autism. *Journal of Autism and Developmental Disorders*. 2007;37(5):929-939; Available from: <https://doi.org/10.1007/s10803-006-0232-9>.
10. Yaneva V, Ha LA, Eraslan S, Yesilada Y, Mitkov R, editors. Detecting autism based on eye-tracking data from web searching tasks. *Proceedings of the 15th International Web for All Conference*; 2018; Available from: <https://doi.org/10.1145/3192714.3192819>.
11. Frazier TW, Klingemier EW, Beukemann M, Speer L, Markowitz L, Parikh S, et al. Development of an objective autism risk index using remote eye tracking. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2016;55(4):301-309; Available from: <https://doi.org/10.1016/j.jaac.2016.01.011>.
12. Duan H, Zhai G, Min X, Che Z, Fang Y, Yang X, et al., editors. A dataset of eye movements for the children with autism spectrum disorder. *Proceedings of the 10th ACM Multimedia Systems Conference*; 2019; Available from: <https://doi.org/10.1145/3304109.3325818>.
13. Wall DP, Kosmicki J, Deluca T, Harstad E, Fusaro VA. Use of machine learning to shorten observation-based screening and diagnosis of autism. *Translational psychiatry*. 2012;2(4):e100-e100; Available from: <https://doi.org/10.1038/tp.2012.10>.
14. Kosmicki J, Sochat V, Duda M, Wall D. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. *Translational psychiatry*. 2015;5(2):e514-e514; Available from: <https://doi.org/10.1038/tp.2015.7>.
15. Duda M, Ma R, Haber N, Wall D. Use of machine learning for behavioral distinction of autism and ADHD. *Translational psychiatry*. 2016;6(2):e732-e732; Available from: <https://doi.org/10.1038/tp.2015.221>.
16. Abbas H, Garberson F, Glover E, Wall DP, editors. Machine learning for early detection of autism (and other conditions) using a parental questionnaire and home video screening. *2017 IEEE International Conference on Big Data (Big Data)*; 2017: IEEE; Available from: <https://doi.org/10.1109/BigData.2017.8258346>.
17. Tiệp VH. *Machine Learning cơ bản*, 2018;.
18. Fuentes J, et al. Autism spectrum disorders. *IACAPAP e-textbook of child and adolescent mental health*. Geneva: International Association for Child and Adolescent Psychiatry and Allied Professions, 2014;.
19. Navon D. Forest before trees: The precedence of global features in visual perception. *Cognitive psychology*. 1977;9(3):353-383. Available from: [https://doi.org/10.1016/0010-0285\(77\)90012-3](https://doi.org/10.1016/0010-0285(77)90012-3).

Analysis of Eye Tracking images to identify children with Autism Spectrum Disorder

Hai Nguyen Ngoc, Ngan Tang Tuyet, Tien Nguyen Thi Thuy, Tin Tran Trung*



Use your smartphone to scan this QR code and download this article

Faculty of Applied Science, Ho Chi Minh City University of Technology, VNUHCM, Vietnam

Correspondence

Tin Tran Trung, Faculty of Applied Science, Ho Chi Minh City University of Technology, VNUHCM, Vietnam

Email: trtrtin@hcmut.edu.vn

History

- Received: 24-6-2021
- Accepted: 04-11-2021
- Published: 25-12-2021

DOI : 10.32508/stdjet.v4i4.862



Copyright

© VNU-HCM Press. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.



ABSTRACT

A number of children with Autism Spectrum Disorder (ASD) are growing to become a big problem in society at the present time. Autism is the neurological disorder that makes the children feel inferior in social communication. Identifying individuals with early autism (under 3 years of age) is difficult and there is still not any medical test for rapid detection of autism. Currently, the clinical diagnosis is mainly based on the observed behavior of the child, and besides, educational and psychological tests are also applied. Eye movements have an important role in an individual's perception and attention in social activities. Non-invasive detection and tracking techniques of the eye movement have been developed over many decades. Nowadays, the Eye Tracking is a technological process that enables the measurement of eye movements, eye positions, and points of gaze. In other words, eye tracking identifies and collects a person's visual data of the eyes in terms of location, objects and duration. It is applied for a variety of different research methods to investigate human behavior. In particular, the Eye Tracking method makes an application for identifying people with Autism Spectrum Disorder (ASD), it has been not only a wide study but also a great interest in recent years. This article provided an overview of children's visual attention with Autism Spectrum Disorder and children with Typical Development (TD). The study has focused on exploring the difference in observed behavior of ASD children and TD children. This study used the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) clustering algorithm for grouping a set of Eye Tracking data, which were collected by the test of children's visual attention to a variety of images. The classification of Eye Tracking data was based on the feature parameters, which were extracted from analyzing data. Moreover, a model for classification and identification of Eye Tracking image data of children groups with ASD and TD was built on Deep Learning algorithms. In this study, the chosen algorithm was Multilayer Perceptron. The children's eye movements dataset was extracted from the paper "A dataset of eye movements for the children with autism spectrum disorder" by Huiyu Duan et al. The results showed that using the Eye Tracking data is highly promising in identifying children with ASD by analyzing the observed behavior of children on images, especially on human images.

Key words: Autism Spectrum Disorder, Eye Tracking, DBSCAN Clustering, Deep Learning, Multilayer Perceptron

Cite this article : Ngoc H N, Tuyet N T, Thuy T N T, Trung T T. **Analysis of Eye Tracking images to identify children with Autism Spectrum Disorder.** *Sci. Tech. Dev. J. – Engineering and Technology*; 4(4):1201-1211.