

Dự Đoán Sự Tập Trung Hướng Đối Tượng của Người Bằng Kỹ Thuật Inverse Reinforcement Learning

19521306 Nguyễn Quang Đại, 19520524 Phan Vỹ Hào, 19522076 Hà Minh Quân,
19520402 Nguyễn Đình Quốc Bảo, và Th.S. Nguyễn Văn Kiệt

Tóm tắt nội dung—Ở đồ án này, nhóm sinh viên tiến hành tái xây dựng lại cấu trúc Inverse Reinforcement Learning sử dụng cho bài toán Scanpath Prediction. Bởi đây là một đồ án mang tính kế thừa, nên nhóm sinh viên tập trung vào đánh giá sự chênh lệch giữa các kết quả được thực nghiệm và các ảnh hưởng của các tham số đến kết quả của mô hình và so sánh nó với kết quả của tác giả. Ngoài ra, nhóm sinh viên cũng dựa trên các kết quả dự đoán được từ mô hình để phân tích và đánh giá các trường hợp xấu và tốt của mô hình dự đoán để nhận định được các hướng phát triển tương lai cho mô hình.

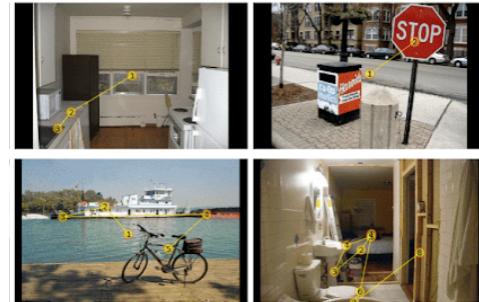
I. GIỚI THIỆU

PHƯƠNG pháp top-down đối với bài toán tìm kiếm hình ảnh thường hướng đến việc xử lý những thông tin top-down như khung cảnh, quan hệ ngữ nghĩa giữa các vật thể,... Ví dụ như khi nhận một tó phở từ nhân viên phục vụ, điều đầu tiên bạn chú ý đến đó là muỗng và đũa, vì nó đóng vai trò vô cùng quan trọng đối với việc thưởng thức ẩm thực.

Đây cũng chính là mục tiêu mà bài báo hướng đến. Nhóm tác giả hướng đến dự đoán đường chuyển động mắt (fixation patterns) trong bài toán tìm kiếm hình ảnh, mô phỏng cách một người tìm kiếm hình ảnh thực tế như thế nào. Nhóm tác giả sử dụng thuật toán Inverse Reinforcement Learning (IRL), tận dụng được ưu điểm của hệ thuật toán bắt chước (Immitation Learning), để học và hiểu những reward function ẩn ý của con người qua nhiều lần quan sát dữ liệu. Ngoài ra, mô hình còn sử dụng phương pháp học đổi ngẫu với hai thành phần Generator và Discriminator, để từ đó học ra những bộ luật có khả năng tạo ra chuyển động mắt tương đương với người. Không những vậy, nhóm tác giả còn tạo ra một bộ dữ liệu mới COCO-Search18, được chọn lọc từ thư viện ảnh COCO gốc.

Đây là một chủ đề thú vị, hấp dẫn và mới lạ đối với nhóm sinh viên. Thứ nhất, nhóm sinh viên chưa từng có kinh nghiệm trong việc sử dụng và huấn luyện mô hình Reinforcement Learning (RL). Thứ hai, bài báo đặt ra nhiều thử thách khi code nhiều lỗi, và không cung cấp code để đánh giá mô hình.

Tại phần II, nhóm sinh viên trình bày quá trình chọn lọc và tạo ra bộ dữ liệu COCO-Search18 chất lượng. Tại phần III, nhóm sẽ trình bày ngắn gọn và tổng quan về kiến trúc mô hình của tác giả. Tại phần IV, nhóm sinh viên sẽ trình bày cách nhóm đã cài đặt, huấn luyện mô hình, cũng như đánh giá về kết quả so với bài báo gốc. Phần cuối V, nhóm sẽ đánh giá bản thân về những thuận lợi, những điều chưa làm được trong đồ án lần này.



Hình 1. Bộ dữ liệu COCO-Search18.

II. BỘ DỮ LIỆU COCO-SEARCH18

Bộ dữ liệu COCO-Search18 là một bộ dữ liệu được tạo ra tại phòng thí nghiệm thuộc dự án *MIT/Tuebingen Saliency Benchmark*. Bộ dữ liệu ghi lại chuyển động mắt (**fixation**) của 10 người trong việc tìm kiếm 18 nhóm đồ vật trong 6,202 bức ảnh chụp từ đời sống và tự nhiên. Tổng cộng nó tạo ra khoảng 300 nghìn cách mắt tìm kiếm. Một nửa số ảnh là có đối tượng cần tìm (target-present - TP), và một nửa số ảnh là không có hình ảnh của đối tượng cần tìm kiếm (target-absent - TA). Một số hình ảnh về bộ dữ liệu ở hình 1.

Những ảnh TP được chọn từ bộ COCO gốc theo 5 tiêu chí: (1) Không dùng ảnh vật nuôi hoặc con người (tránh những thiên vị (bias) có thể có đối với sự tập trung ở người), (2) Hình ảnh chỉ tồn tại duy nhất một vật thể của đối tượng mục tiêu, (3) Kích cỡ của đối tượng mục tiêu không lớn quá 10% kích thước của tổng thể ảnh, (4) Đối tượng mục tiêu không nằm vào trung tâm ảnh, (5) tỉ lệ độ rộng : độ dài của ảnh nằm vào khoảng từ 1.2 đến 2.0.

Những ảnh TA được chọn theo tiêu chí: (1) Không có bất kỳ vật thể nào của đối tượng mục tiêu, (2) Ảnh phải tồn tại ít nhất 2 vật thể liên quan với đối tượng mục tiêu theo định nghĩa của COCO. Ví dụ, lò vi sóng có liên quan với tủ lạnh, lò nướng,... vì cùng thuộc nhóm "gia dụng".

Một nhóm 10 sinh viên bao gồm 6 nam và 4 nữ, tuổi từ 18 - 30, có thị lực bình thường, được yêu cầu đeo một thiết bị ghi lại chuyển động mắt và nhìn tất cả 6,202 hình ảnh. Mỗi người sẽ tham dự 6 phiên (2 giờ/phiên), với mỗi phiên là 1000 ảnh chia đều cho hai nhóm TP và TA. Họ sẽ tìm kiếm một nhóm đối tượng mục tiêu trước khi sang nhóm mục tiêu tiếp theo và đưa ra quyết định ảnh thuộc nhóm "TP" hay "TA" bằng một

thiết bị cầm tay. Tổng cộng nhóm người này đã tạo ra 299,037 chuyển động mắt.

III. MÔ HÌNH

Ngoài dữ liệu, việc đào tạo bất cứ mô hình RL nào luôn cần: (1) định nghĩa các **State** (trạng thái), (2) định nghĩa các **action** (hành động) mà mô hình có thể di chuyển, và (3) cách thức thưởng phạt (**reward policy**) dành cho các cặp state-action mà mô hình tạo ra.

Nhóm tác giả bài báo xem một người tìm kiếm đồ vật trên ảnh là một mô hình hướng đối tượng lý tưởng, với chuyển động mắt của họ (fixation) là một chuỗi của quá trình đưa ra quyết định bởi mô hình. Nói về **action**, tại mỗi bước (time step), mô hình có thể di chuyển đến một vị trí bất kỳ trong ảnh và sẽ nhận một phiên bản hình ảnh bị biến đổi mô phỏng võng mạc người. Hình ảnh ấy có nghĩa là tại vị trí mô hình chọn, hình ảnh sẽ có chất lượng cao (không mờ) (nơi võng mạc mô phỏng chú ý vào), và những vị trí khác sẽ có chất lượng thấp (bị mờ) (mô phỏng cho vùng nằm ngoài vùng trung tâm võng mạc). Các **state** của mô hình được xác định bởi chuỗi điểm ảnh chất lượng cao tạo ra bởi các chuyển động mắt, quá trình biểu diễn trạng thái này sẽ được giải thích rõ hơn ở phần III-B.

Mục tiêu của mô hình là tối đa hóa phần thưởng mà nó nhận được từ việc thay đổi điểm nhìn của mắt. Với mô hình IRL này, việc xác định phần thưởng sử dụng một hàm đánh giá bằng kĩ thuật bắt chước (imitation policy), hàm này được giải thích kỹ hơn ở phần III-C.

A. Tổng quan Inverse Reinforcement Learning (IRL)

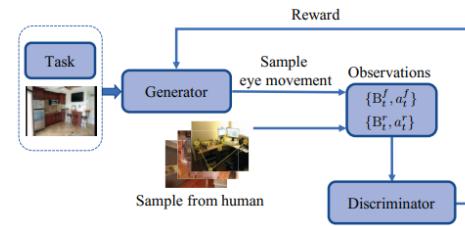
Inverse Reinforcement Learning là một kỹ thuật nâng cao về AI nổi lên trong vài năm gần đây. IRL sinh ra nhằm mục đích giải quyết vấn đề về xung đột ngược trong mô hình RL truyền thống. Trong phương thức hoạt động của các mô hình RL truyền thống, mục tiêu là tìm hiểu quy trình ra quyết định để tạo ra hành vi nhằm tối đa hóa hàm khen thưởng (reward function) được xác định trước. Đối với IRL sẽ đảo ngược lối tư duy của phương thức truyền thống, thay vào đó mô hình sẽ cố gắng trích xuất và định hình hàm khen thưởng (reward function) từ hành vi quan sát được.

IRL chỉ ra rằng chúng ta có thể có được một hàm phần thưởng (reward function) tốt nhất sao cho bám sát với chính sách khen thưởng bằng cách sử dụng các phương pháp lập trình tuyến tính.

Về phương thức hoạt động cơ bản của IRL, toàn này được các nhà nghiên cứu nhận định theo nguyên tắc "học từ con người", bằng cách cấu trúc các phương thức khen thưởng (reward function), mục tiêu tổng quan (agent's objective) và mô hình hóa trạng thái hiện tại ở từng thời điểm dựa theo các quan sát từ dữ liệu. IRL sao chép hành động một con người trong một số nhiệm vụ cụ thể và tìm hiểu những trạng thái nào của môi trường mà con người đang cố gắng đạt được và những mục tiêu cụ thể có thể là gì.

B. Mô hình hóa biểu diễn trạng thái

Với chính sách khen thưởng, nhóm tác giả đề xuất trạng thái Dynamic-Contextual-Belief (DCB) - khả năng tích luỹ thông



Hình 2. Cấu trúc mạng IRL.

tin thông qua các chuyển động mắt. DCB được cấu tạo từ 3 phần:

Fovea: Trung tâm võng mạc của con người và động vật linh trưởng là điểm vàng, ở giữa của vùng điểm vàng có một hố lõm nhỏ được gọi là fovea. Tập trung hơn 50% tế bào thị giác, fovea là nơi chúng ta nhìn màu sắc và chi tiết rõ nhất. Hình ảnh ở ngoài vùng này không rõ chi tiết và mờ dần khi tiến ra xa khỏi nó. Dựa vào cấu trúc đã được đề cập, nhóm tác giả sử dụng hình ảnh gốc với độ phân giải cao tương trưng cho fovea và hình ảnh đã được làm mờ cho vùng ngoài fovea.

Contextual Belief: Theo nhóm tác giả, ngữ cảnh có ảnh hưởng lớn đến quá trình tìm kiếm một vật thể. Ví dụ: Khi chúng ta tìm kiếm bình trà, chúng ta thường sẽ nhìn vào bộ sofa bởi bình trà thường được để trên bàn sofa. Contextual Belief được lấy cảm hứng dựa theo điều này: Mô hình không chỉ được cung cấp các đặc trưng của vật thể mà còn được cung cấp thông tin về các vật thể khác và ngữ cảnh xung quanh của chúng.

Với giả thiết mỗi người có một cách phân tích ngữ cảnh khác nhau (dựa vào kiến thức của người đó), nhóm tác giả cho rằng người đó sẽ tự thiết kế cho mình một bản đồ của họ để tìm kiếm một vật thể trong ngữ cảnh được cung cấp. Từ bản đồ được khởi tạo, thông tin được chuyển tới và hướng dẫn cho fovea để thực hiện việc thu nhận hình ảnh. Sử dụng Panoptic-FPN để tính toán "bản đồ niềm tin" bằng cách cung cấp một ảnh, từ đó sinh ra các pixel-level mask với mỗi nhóm đồ vật. Nhóm tác giả giới thiệu 80 nhóm như thế và 54 nhóm ngữ cảnh của chúng. 134 nhóm này tạo thành bản đồ niềm tin giúp tìm kiếm vật thể.

Dynamics: Để hiểu là sự thay đổi trạng thái sau mỗi lần học. Trạng thái ban đầu được cung cấp dựa trên contextual beliefs của hình ảnh có độ phân giải thấp. Sau mỗi lần học, bản đồ tìm kiếm của hình ảnh phân giải thấp sẽ được thay đổi tương ứng dựa trên bản đồ tìm kiếm của hình ảnh có độ phân giải cao. Mô hình hoá cách thay đổi bằng công thức:

$$B_0 = L \text{ and } B_{t+1} = M_t \odot H + (1 - M_t) \odot B_t \quad (1)$$

Với B_t là trạng thái belief sau t chuyển động mắt, M_t là pixel-level mask sau chuyển động mắt thứ t , L và H lần lượt là bản đồ niềm tin của vật và ngữ cảnh.

C. Chính sách khen thưởng (reward policy)

Nhóm tác giả sử dụng reward function và policy cho bài toán tìm kiếm hình ảnh này bằng thuật toán Generative Adversarial Imitation Learning (GAIL), được mô tả ở ảnh 2. GAIL là một mô hình đối nghịch với một discriminator (thuật toán phân

biệt) và một generator (thuật toán tạo lập). Discriminator có mục tiêu phân biệt thành công những cặp {trạng thái-hành động} thật của người và những cặp giả tạo ra bởi generator. **Policy** mà nhóm tác giả hướng đến là thuật toán generator có khả năng tạo ra những cặp {trạng thái-hành động} giống với hành vi con người nhất.

Gọi D và G là hai mô hình discriminator và generator. E_r là kỳ vọng của phân phối các cặp {trạng thái-hành động} thật của người, còn E_g là kỳ vọng trên phân phối của các cặp tạo ra bởi policy (generator). Hàm mục tiêu của bài toán còn một số hạng thứ ba là norm bình phương kỳ vọng của các gradient, được thêm vào để hội tụ nhanh hơn. Hàm mục tiêu được thể hiện như sau:

$$O = E_r [\log(D(S, a))] + E_f [\log(1 - D(S, a))] - \lambda \times E_r [\|D(S, a)\|^2] \quad (2)$$

Hàm thường của mô hình được định nghĩa dựa trên thuật toán Discriminator:

$$r(S, a) = \log(D(S, a)) \quad (3)$$

Mục tiêu của Generator là một hàm mục tiêu hướng đến xây dựng một bộ luật Reinforcement Learning và tối ưu bằng thuật toán Proximal Policy Optimization:

$$L_\pi = E_\pi [\log(\pi(a | S)) A(S, a)] - E_\pi [\log(\pi(a | S))] \quad (4)$$

Trong công thức 4, hàm thuận lợi A được tính bằng thuật toán Generalized Advantage Estimation.

IV. THỰC NGHIỆM

Dựa trên các kết quả nghiên cứu của nhóm tác giả cite, nhóm sinh viên tiến hành thực nghiệm mô hình đã đề xuất là Scanpath Prediction Framework lên bộ dữ liệu Mscoco. Kế thừa từ nhóm tác giả chính, thang đo đánh giá nhóm sinh viên sử dụng 4 thang đo khác nhau để thực nghiệm thay vì là 5 thang đo như trong nghiên cứu của tác giả chính. Đó là Target Fixation Probability thể hiện mức độ dự đoán chính xác của từng thứ tự chuyển động mắt của kết quả dự đoán, vì vậy thông qua thang đo này có thể đánh giá được hiệu suất của mô hình do nhóm sinh viên thực nghiệm so với kết quả của mô hình gốc do tác giả đề xuất, tiếp theo là Scanpath Ratio có thể đánh giá được khả năng tìm kiếm của mô hình, cùng với đó để đánh giá trên khía cạnh mức độ chính xác dựa trên sự khác nhau của kết quả dự đoán và nhãn được thể hiện qua thang đo Probability Mismatch, và cuối cùng là Multimatch cho việc đánh giá các kết quả trên nhiều khía cạnh như : hình dạng, hướng đi, vị trí, ... của đối tượng được mô hình dự đoán. Bên cạnh đó, nhờ việc tái thiết lập mã nguồn mở do tác giả cung cấp giúp cho nhóm sinh viên có thể so sánh được mức độ chênh lệch giữa kết quả chính và kết quả sau khi được tinh chỉnh các tham số cơ bản, từ đó có thể tiến hành phân tích được những ảnh hưởng của những tham số này đến kết quả. Cuối cùng, nhóm sinh viên dựa trên kết quả dự đoán được tạo ra từ mô hình để thống kê và phân tích các trường hợp cho ra kết quả tốt và xấu để nhận biết được nguyên nhân chính làm cho kết quả tệ đi , nhờ đó nhóm sinh viên có thể hoạch định ra được hướng phát triển cho đề tài vào tương lai.

A. Giới thiệu các mô hình đề xuất

Các mô hình đề xuất được tác giả đưa vào để thể hiện ưu thế của phương pháp IRL trên bài toán Scanpath Prediction. Các mô hình được đề xuất là

- **Random Scanpath** được tác giả xây dựng dựa trên việc sử dụng dự đoán scanpath cho các tập hình ảnh và chọn ngẫu nhiên scanpath được tạo bởi con người cho cùng một mục tiêu tìm kiếm nhưng ở một hình ảnh đầu vào khác
- Tác giả thực nghiệm mô hình **Detector** bằng cách huấn luyện một cấu trúc mạng ConvNet đơn giản dùng để nhận diện vị trí của target và các mẫu tuần tự của ánh nhìn dựa trên confidence map của ảnh.
- Tiếp theo, **Fixation heuristics** được tác giả thực nghiệm cải tiến từ **Detector**, thay thế việc sử dụng confidence map của ảnh để scan đối tượng trong ảnh thì **Fixation heuristics** sử dụng fixation density map từ cấu trúc ConvNet.
- Ké đến là, **BC-CNN** là một phương pháp sao chép hành vi, được xây dựng từ cấu trúc mạng ConvNet để dự đoán ánh nhìn kế tiếp nhờ vào trạng thái DCB đã được mô tả trước đó.
- **BC-LSTM** cũng là một phương pháp sao chép hành vi giống như **BC-CNN** những các trạng thái thể hiện được xử lý và thực hiện với lớp LSTM thay vì phải có quy luật cập nhật trước như **BC-CNN**. **BC-LSTM** đặt ra mục tiêu là học quy tắc cập nhật định kỳ tự động

B. Cài đặt thực nghiệm

Trong quá trình nghiên cứu thuật toán và phương pháp chạy thực nghiệm thì nhóm cũng đã gặp khá nhiều khó khăn, bởi vì mã nguồn mở của nhóm tác giả chính vẫn đang trong quá trình hoàn thiện và code được cung cấp tính toán các kết quả trên thang đo vẫn còn hạn chế, nhưng nhóm sinh viên đã cố gắng để tinh chỉnh và hoàn thiện 4 trên 5 thang đo mà các tác giả chính sử dụng trên môi trường Colab , cụ thể là Colab Pro. Lý do nhóm sinh viên chọn Colab Pro bởi vì GPU mạnh mẽ được cung cấp bởi Google để có thể xử lý được khối lượng tính toán khổng lồ mà mô hình thực hiện trên bộ dữ liệu MS COCO.

Về mặt phương pháp, nhóm dựa trên tiêu chuẩn của mô hình gốc được đề xuất và phát hành trên github với pretrained được sử dụng là Resnet-50-FPN cho Panoptic-FPN để tính contextual beliefs, input ảnh đầu vào sẽ được resize với tỷ lệ 512 : 320 . Ngoài ra các tham số cài đặt mô hình được thiết lập với batch size = 128, con số tinh chỉnh cao hơn của tác giả để tăng tốc độ cập nhật tham số mô hình. Trong quá trình huấn luyện mô hình IRL cho bài toán Scanpath, mô hình sử dụng hai thành phần đó là mạng critic và mạng discriminator . Mạng critic bao gồm 2 lớp tích chập (Convolutional layers) và hai lớp fully-connected layers. Với mạng discriminator sử dụng chung cấu trúc với mạng IRL policy ngoại trừ lớp cuối sử dụng hàm sigmoid.. Toàn bộ quá trình huấn luyện mất khoảng 2 tiếng cho 4500 steps.

Model	TFP-AUC	Probability Mismatch	Scanpath Ratio	Multimatch shape	Multimatch direction	Multimatch length	Multimatch position
Human	5.200	—	0.862	0.490	0.736	0.880	0.910
Random scanpath	0.795	4.407	—	0.295	0.869	0.558	0.849
Detector	4.046	1.166	0.687	0.877	0.676	0.853	0.863
Fixation heuristic	2.154	3.046	0.342	0.873	0.614	0.870	0.850
BC-CNN	3.893	1.328	0.706	0.880	0.669	0.865	0.874
BC-LSTM	1.702	3.497	0.406	0.834	0.567	0.818	0.770
IRL(Nhóm tác giả)	4.509	0.987	0.826	0.886	0.695	0.866	0.885
IRL(Nhóm sinh viên)	4.594	0.918	0.867	0.893	0.706	0.876	0.898

Bảng I

SO SÁNH CÁC KẾT QUẢ THANG ĐO TRÊN NHIỀU MÔ HÌNH KHÁC NHAU TRÊN TẬP DỮ LIỆU TEST CỦA BỘ COCO-SEARCH18

C. Kết quả thực nghiệm

Như kết quả được thể hiện ở bảng I cho thấy được rằng các kết quả từ mô hình IRL tốt hơn nhiều so với các mô hình thực nghiệm khác được tác giả đưa ra, điều này giải thích cho sự vượt trội của mô hình Inverse Reinforcement Learning cho bài toán Scanpath Prediction. Dựa trên thông tin này, nhóm sinh viên đã thực nghiệm và tái tạo lại bộ kết quả của phương pháp IRL trên bài toán và nhận thấy được sự chênh lệch nhỏ giữa 2 kết quả. Sau khi tái đánh giá nhóm sinh viên nhận thấy kết quả được nhóm sinh viên thực hiện nhỉnh hơn một chút ở các thang đo đánh giá **TFP-AUC**, **Multimatch**, **Scanpath Ratio** và **Probability Mismatch** (thang đo mismatch càng thấp càng tốt).

D. Đánh giá

Nhóm sinh viên đã cài đặt thêm code để thực hiện đánh giá mô hình, bổ sung thêm vào mã nguồn mà tác giả đã cung cấp:

- **MultiMatch:** đây là phương pháp sử dụng véc tơ đa chiều để tính toán độ tương đồng của scanpath, lần đầu tiên được đề xuất bởi Jarodzka, Holmqvist và Nyström vào năm 2010. Ở đồ án này chúng tôi sử dụng phương pháp MultiMatch để tính toán bốn phương diện sau của sự tương đồng:
 - **Shape** (hình dạng): được tính bằng cách lấy hiệu hai véc tơ tạo thành từ hai scanpath. Số đo này thể hiện sự khác biệt trong hình dáng của scanpath.
 - **Length** (độ dài) được tính bằng cách lấy hiệu độ dài của hai véc tơ trên trục ngang x.
 - **Direction** (hướng): là số đo khoảng cách góc tính theo hệ số pi, tạo bởi hai véc tơ, nếu hướng của chúng khác thì góc sẽ lớn hơn 0.
 - **Position** (vị trí): đây là sự khác biệt về khoảng cách giữa vị trí của hai véc tơ, được tính bằng khoảng cách Euclide.
- **Target Fixation Probability_AUC:** Tổng tỉ lệ scanpath có số chuyển động mắt (fixation) $\leq n$ (với $0 \leq n \leq 6$)
- **Probability Mismatch:** Độ đo này tính bằng tổng của tất cả những giá trị tuyệt đối từ phép hiệu giữa số đo AUC của người và AUC của máy.
- **Scanpath ratio:** Được tính bằng cách lấy khoảng cách Euclid từ fixation đầu tiên tới tâm của bounding box chia cho tổng độ dài đường đi giữa các fixations (Không tính fixation đầu tiên và cuối cùng).

MultiMatch	Min	Mean	Max	25%	50%	75%
shape	0.670	0.894	0.991	0.864	0.902	0.931
length	0.147	0.707	0.989	0.594	0.721	0.832
direction	0.513	0.877	0.989	0.846	0.889	0.919
position	0.486	0.898	0.991	0.861	0.926	0.953

Bảng II
KẾT QUẢ THANG ĐO MULTIMATCH TRÊN TOÀN TẬP KIỂM THỬ.

Scanpath Ratio	Min	Mean	Max	25%	50%	75%
scanpath	0.081	0.868	1.000	0.874	0.980	1.000

Bảng III
KẾT QUẢ THANG ĐO SCANPATH RATIO TRÊN TOÀN TẬP KIỂM THỬ.

1) *Phân tích kết quả độ đo Multimatch trên toàn tập kiểm thử:* Bảng III cho ta biết kết quả do mô hình của nhóm dự đoán được trên toàn tập kiểm thử và dùng thang đo *MultiMatch* để đánh giá. Từ đây ta thấy mô hình mà nhóm sinh viên cài đặt hoạt động tốt và cho kết quả khá cao ở tất cả 4 tiêu chí đánh giá.

Tiêu chí độ dài (length) có kết quả tệ nhất, và rất nhiều ngoại lệ (outlier), với điểm trung bình vào khoảng 70% nhưng có giá trị nhỏ nhất ở mức 14.7% và bách phân vị 25 rơi vào 59.4%, mặc dù kết quả cao nhất lên tới 98.9%.

Tiêu chí tốt hơn là hướng (direction). Giá trị nhỏ nhất, trung bình, và cao nhất lần lượt là 51.3%, 87.7%, và 98.9%. Bách phân vị 25 và bách phân vị 75 lần lượt ở mức 84.6% và 91.9% cho thấy tiêu chí có kết quả cao, ít lệch và ít có các điểm ngoại lệ.

Hai tiêu chí tốt nhất là hình dáng (shape) và vị trí (position). Giá trị thấp nhất của hai tiêu chí lần lượt là 67% và 48.6%. Giá trị trung bình và cao nhất của hai tiêu chí tương đồng nhau ở mức 89% và 99.1%. Các giá trị bách phân vị cũng cho thấy hai tiêu chí có sự tập trung lớn ở mức cao và ít có ngoại lệ.

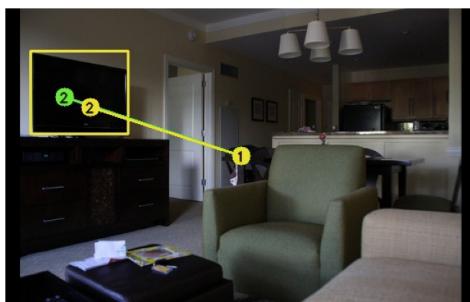


Hình 3. Scanpath từ model có hình dạng tốt nhất toàn tập kiểm thử.



Hình 4. Scanpath từ model có hình dạng tê nhât toàn tập kiểm thử.

Hình 11 và hình 12 cho ta hai trường hợp điển hình tốt và tệ nhất khi đánh giá bằng tiêu chí hình dạng. Trong ảnh thì đường màu xanh lá cây là scanpath của con người, scanpath màu vàng là scanpath được dự đoán từ mô hình. Ở trường hợp tốt nhất, mô hình dự đoán đúng một đường thẳng, giống với con người. Ở trường hợp tệ nhất, mô hình dự đoán đúng một đường thẳng từ điểm ban đầu đến vật thể, khác xa với cách nhìn của con người. Theo sinh viên, do mô hình học tốt và có thể đề xuất được một chuyển động mắt từ điểm đầu tới thẳng vật thể mục tiêu. Tuy chính xác nhưng nó lại không thể hiện được đúng cách con người nhìn. Đây cũng là một điểm mâu thuẫn khiến bài toán khó khăn.



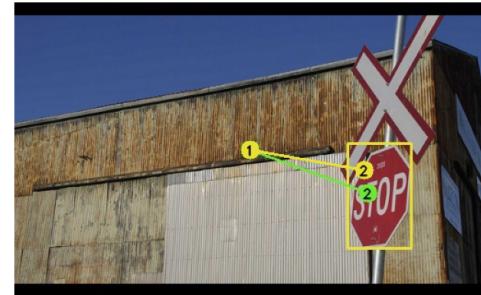
Hình 5. Scanpath từ model có chiều dài khốp nhât toàn tập kiểm thử.



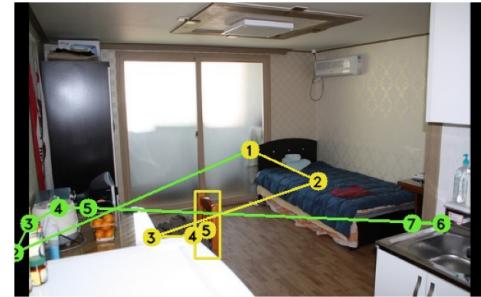
Hình 6. Scanpath từ model có chiều dài tệ nhât toàn tập kiểm thử.

Hình 5 và 6 là hai trường hợp tốt nhất và tệ nhất về tiêu chí độ dài. Theo quan sát của sinh viên, những scanpath tệ về tiêu chí độ dài, đa phần là do ảnh chứa nhiều vật thể nhiễu, khiến mô hình phải dự đoán nhiều điểm chuyển động mắt hơn so với thực tế.

Hình 7 và 8 là trường hợp tốt và tệ nhất về hướng chuyển động mắt. Nếu mô hình đề xuất hướng chuyển động mắt tương



Hình 7. Scanpath từ model có hướng tốt nhât toàn tập kiểm thử.



Hình 8. Scanpath từ model có hướng tệ nhât toàn tập kiểm thử.

tự với con người thì độ đo này sẽ gần 1. Những trường hợp như ảnh 8, vật thể mục tiêu là chiếc ghế, nhưng chính chuyển động mắt của người là chuyển động sai, rối rắm, nên khiến cho mô hình có kết quả tệ.



Hình 9. Scanpath từ model có vị trí tốt nhât toàn tập kiểm thử.



Hình 10. Scanpath từ model có vị trí tệ nhât toàn tập kiểm thử.

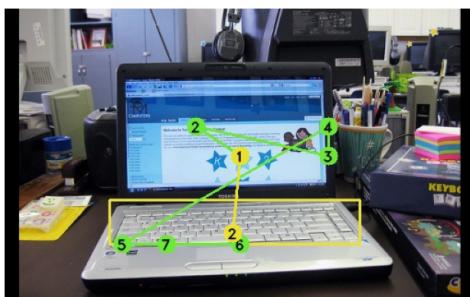
Theo quan sát của nhóm sinh viên, những khung gian có nhiều vật thể nhiễu như ảnh 10 (vật mục tiêu là xe hơi), thì

dễ dàng khiến cho chuyển động mắt mà mô hình đề xuất bị sai.

2) *Phân tích kết quả độ đo Scanpath trên toàn tập kiểm thử:* Kết quả mô hình sử dụng thang đo *Scanpath Ratio* trên toàn tập kiểm thử ở bảng III cho thấy mô hình được cài đặt có kết quả khá tốt. Với kết quả tốt nhất đạt 100%, nhỏ nhất đạt 8.1%, ta thấy với thang đo *Scanpath Ratio* tỉ lệ kết quả trải dài, một số cho kết quả tối đa. Các bách phân vị cũng cho thấy sự phân bố lớn ở mức cao (>0.87%).



Hình 11. Scanpath từ model có hình dạng tốt nhất toàn tập kiểm thử.



Hình 12. Scanpath từ model có hình dạng tệ nhất toàn tập kiểm thử.

Hình 11 và hình 12 cho ta hai trường hợp điển hình tốt và tệ nhất khi đánh giá bằng thang đo *Scanpath Ratio*. Tương tự với thang đo *Multimatch*, đường màu xanh lá cây là *scanpath* của con người, *scanpath* màu vàng là *scanpath* được dự đoán từ mô hình, ta thấy ở trường hợp tốt nhất, mô hình dự đoán đúng một đường thẳng trùng với con người; trong khi đó với trường hợp tệ nhất, mô hình chỉ dự đoán một đường thẳng từ điểm ban đầu đến vật thể, trong khi theo cách nhìn của con người thì rất phức tạp. Theo sinh viên, do mô hình học tốt và có thể đề xuất được một chuyển động mắt từ điểm đầu tới thẳng vật thể mục tiêu. Tuy chính xác nhưng nó lại không thể hiện được đúng cách con người nhìn. Đây cũng là một điểm mâu thuẫn khiến bài toán khó khăn.

3) :

V. KẾT LUẬN

Kết quả đạt được: Nhóm sinh viên đã nghiên cứu và cài đặt thành công mô hình Inverse Reinforcement Learning với kết quả bằng hoặc cao hơn nhóm tác giả của bài báo gốc. Bên cạnh đó, nhóm sinh viên còn bổ sung thêm mã nguồn của các độ đo để đánh giá kết quả trên tập test và trực quan hóa kết quả dự đoán, đây là điều nhóm sinh viên đã làm được nhiều

hơn bài báo. Cuối cùng, nhóm sinh viên tiến hành phân tích kết quả dự đoán trên tập test, và đánh giá những trường hợp tiêu biểu nhất, dễ lối nhất của mô hình. Kết quả chung cho thấy nhóm sinh viên đạt được kết quả rất ấn tượng.

Những điểm hạn chế: Tác giả bài báo đã đề xuất phương pháp đánh giá mô hình bằng 5 thuật toán Deep Learning khác nhau. Tuy nhiên tác giả không cung cấp mã nguồn, với thời gian ngắn trong đồ án, nhóm sinh viên chưa thể nghiên cứu và triển khai thêm 4 thuật toán tương đương với tác giả.

LỜI CẢM ƠN

Nhóm sinh viên bày tỏ lòng biết ơn sâu sắc đến Thạc sĩ Nguyễn Văn Kiệt, người đã đồng hành cùng nhóm trong suốt hành trình nghiên cứu này với góc nhìn chuyên môn sắc sảo cùng những lời khuyên quý báu.

TÀI LIỆU

- [1] Yang, Z., Huang, L., Chen, Y., Wei, Z., Ahn, S., Zelinsky, G., ... & Hoai, M. (2020). Predicting goal-directed human attention using inverse reinforcement learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 193-202).
- [2] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755). Springer, Cham.
- [3] Zelinsky, G., Yang, Z., Huang, L., Chen, Y., Ahn, S., Wei, Z., ... & Hoai, M. (2019). Benchmarking gaze prediction for categorical visual search. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 0-0).
- [4] Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2015). A comparison of scanpath comparison methods. Behavior research methods, 47(4), 1377-1392.
- [5] Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. Behavior research methods, 44(4), 1079-1100.
- [6] Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. Numerische Mathematik, 1(1), 269 - 271. <https://doi.org/10.1007/BF01386390>
- [7] Jarodzka, H., Holmqvist, K. & Nyström, M. (2010). A vector-based, multidimensional scanpath similarity measure. In ETRA '02: Proceedings of the 2010 symposium on eye tracking research and applications, ACM, New York. <https://doi.org/10.1145/1743666.1743718>