

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC ĐÀ LẠT**



BÁO CÁO TỔNG KẾT

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO VÀO PHÁT TRIỂN GAME HỌC NGOẠI NGỮ

Thuộc nhóm ngành khoa học: Công nghệ thông tin

Đà Lạt, ngày 24 tháng 4 năm 2024

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC ĐÀ LẠT**



BÁO CÁO TỔNG KẾT

**ỨNG DỤNG TRÍ TUỆ NHÂN TẠO VÀO PHÁT
TRIỂN GAME HỌC NGOẠI NGỮ**

Mã số đề tài

Thuộc nhóm ngành khoa học: Công nghệ thông tin

Đà Lạt, ngày 24 tháng 4 năm 2024

MỤC LỤC

DANH MỤC BẢNG, BIỂU ĐỒ

DANH SÁCH SINH VIÊN THAM GIA ĐỀ TÀI NGHIÊN CỨU

MỞ ĐẦU	1
1. Lý do chọn đề tài	1
2. Tổng quan tình hình nghiên cứu	1
3. Mục tiêu nghiên cứu	2
4. Cách tiếp cận, phương pháp nghiên cứu, phạm vi nghiên cứu	2
4.1 Cách tiếp cận	2
4.2 Phương pháp nghiên cứu	2
4.3 Phạm vi nghiên cứu	2
5. Đóng góp mới của đề tài	5
6. Kết cấu báo cáo nghiên cứu	5
CHƯƠNG 1: CƠ SỞ LÝ THUYẾT VỀ TRÍ TUỆ NHÂN TẠO	6
1.1. Trí tuệ nhân tạo (AI)	6
1.1.1 Định nghĩa	6
1.1.2. Các lĩnh vực ứng dụng của AI	6
1.1.3. Ưu điểm và hạn chế của AI	6
1.2. Computer Vision (Thị giác máy tính)	7
1.3. Các ứng dụng của AI trong giáo dục	8
1.3.1. Cá nhân hóa việc học tập	8
1.3.2. Hỗ trợ giáo viên	8
1.3.3. Hỗ trợ đánh giá bài làm của học sinh	8
1.3.4. Hỗ trợ phát hiện gian lận	9
1.3.5. Hỗ trợ quản lý học sinh	9

1.3.6. Dịch thuật	9
CHƯƠNG 2: CÁC CÔNG NGHỆ SỬ DỤNG	10
2.1. Unity	10
2.1.1. Khái niệm	10
2.1.2. Chức năng chính	10
2.1.3. Ưu điểm	10
2.1.4. Nhược điểm	11
2.2. OpenCV	11
2.2.1. Khái niệm	11
2.2.2. Lịch sử phát triển	12
2.2.3. Ưu điểm	12
2.2.4. Nhược điểm	12
2.2.5. Ứng dụng	12
2.3. MediaPipe	13
2.3.1. Khái niệm	13
2.3.2. Lịch sử phát triển	13
2.3.3. Ưu điểm	14
2.3.4. Nhược điểm	14
2.3.5. Ứng dụng	14
2.4. CVZone	14
2.4.1. Khái niệm	15
2.4.2. Chức năng chính	15
2.4.3. Ưu điểm	15
2.4.4. Nhược điểm	15
2.4.5. Ứng dụng	16
CHƯƠNG 3: ỨNG DỤNG GAME HỌC NGOẠI NGỮ	17
3.1. Thu thập dữ liệu từ vựng	17

3.2. Dò tìm bàn tay (Hand tracking) trong CV Zone	17
3.3. Kết nối CV Zone với Unity	19
3.3.1. Gửi dữ liệu từ CV Zone	19
3.3.2. Nhận dữ liệu trong Unity	19
3.4. Ứng dụng Game học ngoại ngữ.....	20
KẾT LUẬN	23
DANH MỤC TÀI LIỆU THAM KHẢO	24

DANH MỤC BẢNG, BIỂU ĐỒ

STT	NỘI DUNG		TRANG
01	Bảng 1	Bảng thông tin số từ vựng được sử dụng theo cấp bậc của CEFR	

DANH SÁCH SINH VIÊN THAM GIA ĐỀ TÀI NGHIÊN CỨU

Mã số sinh viên	Họ và tên	Khoa – Lớp	Nhiệm vụ trong đề tài
2111846	Lê Trần Anh Khôi	Khoa Công nghệ thông tin – CTK45A	- Tìm hiểu và khảo sát ứng dụng học tiếng Anh liên quan - Tìm hiểu cách thức xây dựng ứng dụng Game - Xây dựng ứng dụng
2115208	Lê Trung Hòa Hiếu	Khoa Công nghệ thông tin – CTK45A	- Tìm hiểu về Unity - Thu thập tập dữ liệu về từ vựng tiếng Anh - Kiểm tra ứng dụng khi hoàn thành
2212343	Triệu Quang Học	Khoa Công nghệ thông tin – CTK46B	- Tìm hiểu về ứng dụng của trí tuệ nhân tạo trong game - Thu thập dữ liệu về từ vựng tiếng Anh - Xây dựng ứng dụng
2212343	Đinh Lâm Gia Bảo	Khoa Công nghệ thông tin – CTK46B	- Tìm hiểu về AI trong Unity - Lên ý tưởng cho Game học ngoại ngữ - Viết báo cáo
2212456	Lê Thành Thái	Khoa Công nghệ thông tin – CTK46B	- Lên ý tưởng cho Game học ngoại ngữ - Xây dựng ứng dụng - Viết báo cáo

MỞ ĐẦU

1. Lý do chọn đề tài

Thế giới ảo 3D đang trở thành một xu hướng mới, mở ra cánh cửa đến những trải nghiệm mới mẻ và đầy tiềm năng. Tuy nhiên, sự tiếp cận với thế giới ảo này còn hạn chế do rào cản về chi phí với những thiết bị đắt đỏ, chưa phổ biến hạn chế khả năng tiếp cận của một số người nhất định.

Nhận thức được điều này, nhóm chúng tôi quyết tâm mang thế giới ảo 3D đến gần hơn với mọi người. Sử dụng webcam - thiết bị sẵn có trên hầu hết máy tính bàn, laptop, chúng tôi phát triển trò chơi học tiếng Anh ứng dụng công nghệ hand tracking.

2. Tổng quan tình hình nghiên cứu

Trong bối cảnh toàn cầu hóa và sự tăng cường hợp tác quốc tế, việc học ngoại ngữ trở thành một yêu cầu vô cùng quan trọng và cần thiết. Có rất nhiều phương pháp, cách tiếp cận để học một ngoại ngữ như: tham gia các lớp học truyền thống, học qua sách giáo trình, học trực tuyến hay sử dụng ứng dụng di động... một phương pháp trong số đó đã thu hút được sự quan tâm của các nhà nghiên cứu trong lĩnh vực giáo dục chính là phương pháp học thông qua các trò chơi. Phương pháp này đã được nhiều nghiên cứu chứng minh là hiệu quả, giúp dễ học và dễ tiếp thu [1]. Với những trò chơi máy tính nghiêm túc, phương pháp học thông qua chơi trò chơi cũng có thể nâng cao nhận thức của học sinh về quy trình giảng dạy và chuyển hóa kiến thức một cách hiệu quả [2]. Cùng với sự bùng nổ của cách mạng công nghiệp lần thứ tư, trí tuệ nhân tạo (AI) đã và đang được ứng dụng trong nhiều lĩnh vực của đời sống, trong đó có ngành công nghiệp Game. Lập trình trò chơi kết hợp với AI không chỉ tạo cho trò chơi có tính sinh động hơn, mang lại cảm xúc nhiều hơn mà còn giúp trò chơi thông minh hơn, tự xử lý được nhiều vấn đề để tương tác tốt hơn với người chơi. Với việc đưa AI vào trong lập trình game, các nhà phát triển cố gắng mang lại trải nghiệm tương tác có giá trị cho người chơi. Những trải nghiệm này bắt nguồn từ hiệu ứng tích lũy của một số yếu tố trò chơi được phối hợp, bao gồm đồ họa, âm thanh, thời gian chơi, lời kể, thử thách và nội dung tương tác trực tiếp với người chơi (đồng minh, đối thủ hoặc các đối

tượng khác). Do đó, AI đang trở thành một công cụ đắc lực để có thể giúp các nhà thiết kế và phát triển xử lý được các vấn đề phức tạp của quá trình mà người chơi thực hiện trong Game [3]. Với mong muốn áp dụng các công nghệ AI vào để học tiếng Anh hiệu quả, nhóm chúng tôi sẽ sử dụng phần mềm Unity để tạo ra một trò chơi có sự tham gia của trí tuệ nhân tạo để giúp cải thiện, đổi mới việc học một ngoại ngữ một cách tốt hơn, đồng thời giúp người học tương tác tốt và hiệu quả trong quá trình vừa học vừa chơi của họ.

3. Mục tiêu nghiên cứu

- Tìm hiểu lý thuyết lập trình game với công cụ Unity. Đây là một công cụ phổ biến và mạnh mẽ trong việc thiết kế game và được nhiều nhà phát triển lựa chọn
- Tìm hiểu ứng dụng AI trong lập trình Game thông qua thư viện Python CV Zone.
- Xây dựng một trò chơi giúp học và rèn luyện tiếng Anh, có sử dụng trí tuệ nhân tạo để tăng khả năng tương tác và xử lý thông tin với người chơi.

4. Cách tiếp cận, phương pháp nghiên cứu, phạm vi nghiên cứu

4.1 Cách tiếp cận

Đề tài khảo sát chức năng của các ứng dụng học ngoại ngữ hiện nay, từ đó áp dụng các kỹ thuật lập trình để xây dựng ứng dụng trò chơi với Unity.

4.2 Phương pháp nghiên cứu

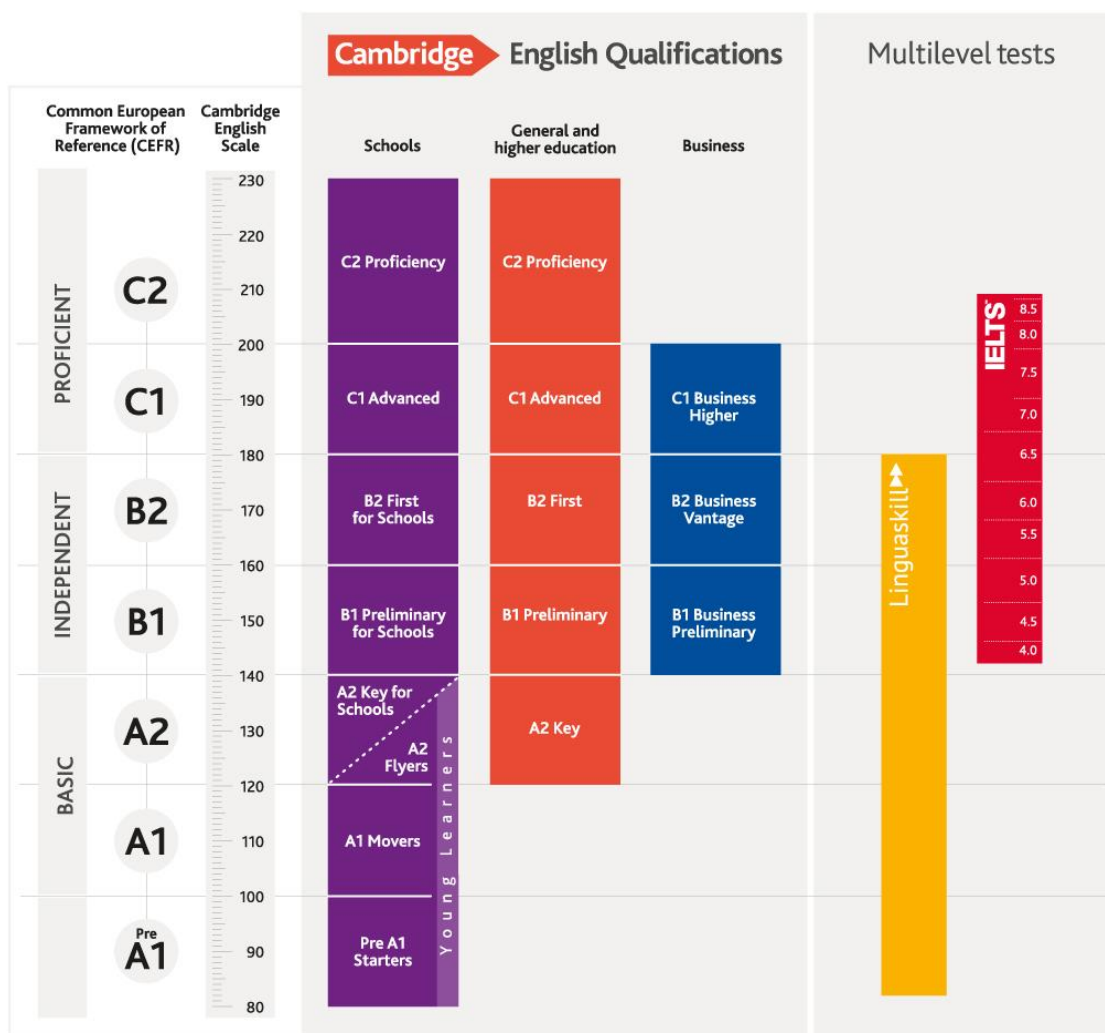
Đề tài áp dụng phương pháp tìm hiểu và thực nghiệm nhằm đánh giá và so sánh hiệu quả của việc áp dụng và không áp dụng trí tuệ nhân tạo vào việc xây dựng ứng dụng trò chơi phục vụ cho mục đích học tập.

4.3 Phạm vi nghiên cứu

Đề tài chỉ tập trung nghiên cứu, ứng dụng máy học vào xây dựng một trò chơi theo lượt để học hoặc rèn luyện kiến thức liên quan đến từ vựng tiếng Anh theo các cấp của chứng chỉ CEFR.

Chứng chỉ CEFR:

Khung Tham Chiếu Ngôn Ngữ Chung Châu Âu (CEFR) là một tiêu chuẩn quốc tế để mô tả khả năng sử dụng ngôn ngữ. Nó mô tả khả năng sử dụng ngôn ngữ theo 6 bậc, từ cấp độ A1 cho người mới bắt đầu, cho đến cấp độ C2 cho những người đã thành thạo một ngôn ngữ. Điều này giúp cho bất kỳ ai tham gia vào việc giảng dạy và kiểm tra đánh giá ngôn ngữ, chẳng hạn như giáo viên hoặc người học, dễ dàng thấy được trình độ của các cấp độ khác nhau. Điều đó cũng có nghĩa là các nhà tuyển dụng và các tổ chức giáo dục có thể dễ dàng so sánh các bằng cấp của chúng tôi với các kỳ thi khác ở quốc gia của họ. (CAMBRIDGE, n.d.)



5. Đóng góp mới của đề tài

Về mặt lý luận: Kết quả nghiên cứu góp phần làm cơ sở cho các nghiên cứu tiếp theo có sử dụng thư viện CV Zone kết hợp với Unity Engine.

Về mặt thực tiễn: Đề tài sẽ giúp mọi người thấy được khả năng ứng dụng của các thư viện của Python, cụ thể là các thư viện nhận diện hình ảnh như CV Zone vào các công nghệ gần gũi với cuộc sống như camera, máy ảnh và tạo ra các phần mềm hữu ích.

6. Kết cấu báo cáo nghiên cứu

Ngoài phần Mở đầu và Kết luận, Danh mục bảng biểu, Danh mục tài liệu tham khảo, báo cáo đề tài dự kiến kết cấu gồm 3 Chương:

Chương 1: Cơ sở lý thuyết về Trí tuệ nhận tạo.

Chương 2: Các công nghệ sử dụng.

Chương 3: Ứng dụng Game học ngoại ngữ.

CHƯƠNG 1: CƠ SỞ LÝ THUYẾT VỀ TRÍ TUỆ NHÂN TẠO

1.1. Trí tuệ nhân tạo (AI)

1.1.1 Định nghĩa

Trí tuệ nhân tạo là một nhánh của khoa học và công nghệ tạo ra các máy móc và chương trình máy tính thông minh để thực hiện các nhiệm vụ khác nhau đòi hỏi trí tuệ của con người. (Kaplan, 2021)

1.1.2. Các lĩnh vực ứng dụng của AI

Y tế: Chẩn đoán bệnh, phát triển thuốc, phẫu thuật robot,...

Tài chính: Phân tích thị trường, chống gian lận, quản lý rủi ro,...

Giao thông: Xe tự lái, hệ thống quản lý giao thông thông minh,...

Sản xuất: Tự động hóa quy trình sản xuất, dự đoán bảo trì,...

Giáo dục: Cá nhân hóa việc học tập, đánh giá học sinh, hỗ trợ giáo viên,...

Nhiều lĩnh vực khác: Giải trí, bán lẻ, dịch vụ khách hàng,...

1.1.3. Ưu điểm và hạn chế của AI

- Ưu điểm:

Tự động hóa các tác vụ: AI có thể tự động hóa nhiều tác vụ tốn thời gian và công sức do con người thực hiện.

Nâng cao hiệu quả: AI có thể giúp giải quyết các vấn đề một cách nhanh chóng và hiệu quả hơn con người.

Phân tích dữ liệu: AI có thể phân tích lượng dữ liệu khổng lồ mà con người không thể xử lý được.

Đưa ra dự đoán: AI có thể đưa ra dự đoán chính xác về các sự kiện trong tương lai.

Sáng tạo: AI có thể hỗ trợ con người sáng tạo ra những sản phẩm và dịch vụ mới.

- Hạn chế:

Thiếu khả năng sáng tạo: AI không có khả năng sáng tạo như con người, mà chỉ có thể học hỏi và mô phỏng những gì đã được con người tạo ra.

Thiếu khả năng cảm nhận: AI không có khả năng cảm nhận cảm xúc và suy nghĩ như con người.

Nguy cơ thiên vị: AI có thể bị ảnh hưởng bởi dữ liệu được sử dụng để huấn luyện, dẫn đến thiên vị trong việc đưa ra quyết định.

Mất việc làm: AI có thể thay thế một số công việc của con người, dẫn đến nguy cơ mất việc làm.

Nguy cơ lạm dụng: AI có thể bị lạm dụng cho mục đích xấu, gây hại cho con người và xã hội.

1.2. Computer Vision (Thị giác máy tính)

Computer Vision - Thị giác Máy tính là sự kết hợp giữa xử lý hình ảnh và nhận dạng mẫu. Đầu ra của quá trình Thị giác Máy tính là sự hiểu biết về hình ảnh. Sự phát triển của lĩnh vực này được thực hiện bởi thích ứng với khả năng thị giác của con người trong việc tiếp nhận thông tin. Thị giác Máy tính là hướng nghiên cứu về trích xuất thông tin từ hình ảnh, là một phần của Đồ họa máy tính. Sự phát triển của Thị giác Máy tính phụ thuộc vào hệ thống công nghệ máy tính, cho dù chất lượng hình ảnh cải thiện hoặc nhận dạng hình ảnh. Có sự trùng lặp với Xử lý hình ảnh về các kỹ thuật cơ bản, và một số tác giả sử dụng cả hai thuật ngữ thay thế cho nhau. (Victor Wiley, 2018)

Ví dụ:

Robot hỗ trợ phẫu thuật: Hệ thống Thị giác Máy tính được sử dụng trong robot hỗ trợ phẫu thuật để giúp bác sĩ thực hiện các thao tác phẫu thuật chính xác hơn.

Hệ thống tự lái của xe ô tô: Hệ thống Thị giác Máy tính được sử dụng trong hệ thống tự lái của xe ô tô để thu thập dữ liệu không gian xung quanh xe.

Kiểm tra sản phẩm: Hệ thống Thị giác Máy tính được sử dụng trong kiểm tra sản phẩm để phát hiện lỗi sản phẩm.

Ngoài ra, còn có một số loại hình AI phổ biến khác như:

Xử lý ngôn ngữ tự nhiên (Natural Language Processing - NLP): Cho phép máy tính hiểu và xử lý ngôn ngữ con người.

Học tăng cường (Reinforcement Learning): Cho phép máy tính học hỏi thông qua tương tác với môi trường.

Hệ thống chuyên gia (Expert Systems): Sử dụng kiến thức và kinh nghiệm của con người để giải quyết các vấn đề phức tạp.

1.3. Các ứng dụng của AI trong giáo dục

1.3.1. Cá nhân hóa việc học tập

AI có thể phân tích dữ liệu học tập của học sinh như điểm số, bài tập, thời gian học tập,... để đánh giá trình độ, điểm mạnh, điểm yếu và đưa ra phương pháp học tập phù hợp với từng cá nhân. Ví dụ, AI có thể đề xuất các bài tập luyện tập phù hợp với trình độ của học sinh, hoặc tạo ra các bài học cá nhân hóa dựa trên sở thích và nhu cầu của từng người.

1.3.2. Hỗ trợ giáo viên

AI có thể giúp giáo viên soạn bài giảng, chấm điểm bài tập, và giải đáp thắc mắc của học sinh một cách nhanh chóng. Ngoài ra, AI cũng có thể cung cấp cho giáo viên dữ liệu học tập của học sinh, giúp giáo viên đánh giá hiệu quả việc học tập và điều chỉnh phương pháp giảng dạy phù hợp.

1.3.3. Hỗ trợ đánh giá bài làm của học sinh

AI có thể tự động đánh giá các bài tập và bài kiểm tra của học sinh, giúp giáo viên tiết kiệm thời gian và công sức. AI cũng có thể đánh giá khả năng giao tiếp tiếng Anh của học sinh một cách chính xác và khách quan hơn so với đánh giá của giáo viên.

1.3.4. Hỗ trợ phát hiện gian lận

AI có thể phát hiện gian lận trong các bài kiểm tra và bài thi trực tuyến, giúp đảm bảo tính công bằng trong việc đánh giá học sinh. AI cũng có thể phát hiện đạo văn trong các bài luận và bài báo cáo của học sinh, giúp duy trì đạo đức học tập.

1.3.5. Hỗ trợ quản lý học sinh

AI có thể giúp nhà trường quản lý sinh viên, theo dõi việc đi học, điểm danh và đưa ra các báo cáo về tình hình học tập của sinh viên. AI cũng có thể giúp nhà trường quản lý các hoạt động ngoại khóa và các chương trình học tập của sinh viên.

1.3.6. Dịch thuật

AI có thể dịch văn bản và giọng nói từ các ngoại ngữ sang tiếng Việt và ngược lại, giúp học sinh hiểu bài giảng và tài liệu học tập dễ dàng hơn.

CHƯƠNG 2: CÁC CÔNG NGHỆ SỬ DỤNG

2.1. Unity

2.1.1. Khái niệm

Unity là một công cụ phát triển game đa nền tảng mạnh mẽ và phổ biến được sử dụng để tạo ra các trò chơi 2D, 3D, VR và AR. Unity được sử dụng bởi các nhà phát triển game ở mọi cấp độ, từ người mới bắt đầu đến chuyên gia, để tạo ra các trò chơi cho nhiều nền tảng khác nhau, bao gồm PC, Mac, Linux, di động, web và console.

2.1.2. Chức năng chính

- Tạo cảnh trò chơi: Unity cung cấp cho người dùng các công cụ để tạo ra các thế giới trò chơi 2D và 3D chi tiết, sử dụng các mô hình 3D, hình ảnh, âm thanh và video. Unity cũng có các công cụ tạo hiệu ứng đặc biệt như ánh sáng, bóng tối, nước và lửa.

- Lập trình: Unity sử dụng ngôn ngữ lập trình C# để viết mã, kiểm soát hành vi của các đối tượng và hệ thống trò chơi. Lập trình cho phép tạo ra nhân vật, kẻ thù, vật phẩm, menu, hệ thống điểm số, v.v.

- Hệ thống trò chơi: Unity cung cấp các hệ thống trò chơi tích hợp sẵn như hệ thống vật lý, hệ thống âm thanh và hệ thống ánh sáng để tạo ra các trò chơi thực tế và hấp dẫn. Unity cũng cung cấp công cụ để tạo ra các hệ thống trò chơi của riêng người dùng.

- Cửa hàng Unity Asset Store: Cung cấp nhiều tài nguyên miễn phí và trả phí như mô hình 3D, âm thanh và mã, giúp tiết kiệm thời gian và công sức.

- Xuất bản: Unity cho phép xuất bản trò chơi cho nhiều nền tảng như PC, Mac, Linux, di động, web và console, đồng thời cung cấp công cụ quảng bá và tiếp thị trò chơi.

2.1.3. Ưu điểm

Dễ sử dụng: Giao diện người dùng trực quan, dễ sử dụng, giúp người mới bắt đầu dễ dàng học cách sử dụng.

Đa nền tảng: Tạo ra các trò chơi cho nhiều nền tảng, tiếp cận nhiều người chơi hơn.

Cộng đồng lớn: Cộng đồng nhà phát triển game tích cực, sẵn sàng giúp đỡ và hỗ trợ.

Nhiều tài nguyên: Nhiều tài nguyên học tập miễn phí và trả phí giúp học cách sử dụng Unity hiệu quả.

Miễn phí: Phiên bản miễn phí với đầy đủ tính năng, giúp người dùng bắt đầu phát triển game mà không cần trả phí.

2.1.4. Nhược điểm

Unity là công cụ mạnh mẽ với nhiều tính năng, đòi hỏi thời gian để học cách sử dụng.

Hiệu suất: Hiệu suất trò chơi Unity có thể phụ thuộc vào phần cứng máy tính của người chơi.

Kích thước tệp: Trò chơi Unity có thể có kích thước tệp lớn, gây khó khăn trong việc tải xuống và cài đặt.

Hệ thống AI: Unity có tích hợp một số tính năng AI cơ bản, nhưng người dùng có thể cần sử dụng thư viện bên thứ ba để ứng dụng các tính năng AI phức tạp hơn.

2.2. OpenCV

2.2.1. Khái niệm

OpenCV (Open Source Computer Vision Library - Thư viện thị giác máy tính mã nguồn mở) là một thư viện phần mềm Máy học và Thị giác Máy tính mã nguồn mở. OpenCV được xây dựng để cung cấp cơ sở hạ tầng chung cho các ứng dụng Thị giác Máy tính và tăng tốc việc sử dụng Nhận thức máy trong các sản phẩm thương mại. Là

sản phẩm được cấp phép của Apache 2, OpenCV giúp doanh nghiệp dễ dàng sử dụng và sửa đổi mã. (OpenCV, n.d.)

2.2.2. Lịch sử phát triển

OpenCV được phát triển ban đầu bởi Intel vào năm 2000. Sau đó, nó được phát hành theo giấy phép nguồn mở BSD vào năm 2005. Hiện tại, OpenCV được quản lý bởi Open Source Computer Vision Library (OSCV) Community.

2.2.3. Ưu điểm

Mã nguồn mở: Miễn phí cho mục đích thương mại và phi thương mại.

Cộng đồng lớn: Cộng đồng người dùng và nhà phát triển tích cực, sẵn sàng giúp đỡ và hỗ trợ.

Nhiều tài nguyên: Nhiều tài liệu, hướng dẫn và ví dụ học tập có sẵn.

Hỗ trợ nhiều nền tảng: Hỗ trợ nhiều hệ điều hành và ngôn ngữ lập trình.

Hiệu suất cao: Cung cấp các thuật toán tối ưu hóa cho xử lý ảnh và video.

2.2.4. Nhược điểm

Có thể khó học đối với người mới bắt đầu.

Yêu cầu kiến thức lập trình: Yêu cầu hiểu biết về ngôn ngữ lập trình như C++ hoặc Python.

Cập nhật thường xuyên: Cần cập nhật thư viện thường xuyên để sử dụng các tính năng mới nhất.

2.2.5. Ứng dụng

Xử lý ảnh: Phát hiện cạnh, nhận dạng khuôn mặt, phân tích kết cấu, v.v.

Xử lý video: Theo dõi chuyển động, phân tích hành vi, nhận dạng đối tượng, v.v.

Tăng cường thực tế: Ghép ảnh ảo vào thế giới thực.

Robot và tự động hóa: Điều khiển robot và hệ thống tự động bằng thị giác máy tính.

Y tế: Phân tích hình ảnh y tế, chẩn đoán bệnh, v.v.

2.3. MediaPipe

2.3.1. Khái niệm

MediaPipe là một thư viện mã nguồn mở do Google phát triển, được sử dụng cho xử lý ảnh và video trên nhiều nền tảng như di động, web và desktop. Nó cung cấp các công cụ mạnh mẽ để thực hiện các tác vụ như:

Phát hiện khuôn mặt: Xác định vị trí và kích thước của khuôn mặt trong ảnh hoặc video.

Phát hiện cơ thể: Xác định vị trí của các khớp chính trên cơ thể như vai, hông, đầu gối và mắt cá chân.

Theo dõi tay: Theo dõi chuyển động của tay trong ảnh hoặc video, xác định vị trí của các ngón tay.

Nhận dạng cử chỉ: Nhận dạng các cử chỉ tay đơn giản, chẳng hạn như vẫy tay, giơ ngón tay cái và v.v.

Xử lý ảnh: Thực hiện các thao tác xử lý ảnh cơ bản như xoay, thay đổi kích thước, cắt ghép và điều chỉnh độ sáng.

Hiệu ứng video: Thêm hiệu ứng vào video như bộ lọc, chuyển tiếp và văn bản.

(Google, n.d.)

2.3.2 Lịch sử phát triển

MediaPipe được phát triển bởi Google AI vào năm 2019. Nó được phát hành theo giấy phép nguồn mở Apache 2.0.

2.3.3. Ưu điểm

Dễ sử dụng: Cung cấp các API đơn giản và trực quan.

Hiệu suất cao: Sử dụng công nghệ học máy tiên tiến để xử lý ảnh và video hiệu quả.

Mô hình được đào tạo sẵn: Cung cấp nhiều mô hình được đào tạo sẵn cho các tác vụ phổ biến như phát hiện khuôn mặt, phát hiện cơ thể, theo dõi tay, v.v.

Hỗ trợ nhiều nền tảng: Hỗ trợ nhiều hệ điều hành và ngôn ngữ lập trình.

Tích hợp với TensorFlow: Dễ dàng tích hợp với các mô hình học máy TensorFlow.

2.3.4. Nhược điểm

Thư viện mới: Còn đang được phát triển, một số tính năng chưa hoàn thiện.

Cộng đồng nhỏ hơn: Cộng đồng người dùng và nhà phát triển nhỏ hơn so với OpenCV.

Yêu cầu kiến thức học máy: Cần hiểu biết về học máy để sử dụng các tính năng tiên tiến.

2.3.5. Ứng dụng

Giao diện người dùng: Phát hiện cử chỉ, theo dõi chuyển động, v.v.

Thực tế tăng cường: Ghép ảnh ảo vào thế giới thực.

Giáo dục: Theo dõi sự tập trung của học sinh, đánh giá hiệu quả học tập, v.v.

Thể thao: Phân tích chuyển động của vận động viên, theo dõi hiệu suất thi đấu, v.v.

Chăm sóc sức khỏe: Phân tích hình ảnh y tế, hỗ trợ chẩn đoán, v.v.

2.4. CVZone

2.4.1. Khái niệm

CVZone là thư viện Python mã nguồn mở được sử dụng cho xử lý ảnh và video bằng OpenCV.

2.4.2. Chức năng chính

CVZone cung cấp các lớp và hàm đơn giản để thực hiện các tác vụ phổ biến như:

- Phát hiện khuôn mặt: Xác định vị trí và kích thước của khuôn mặt trong ảnh hoặc video.
- Theo dõi khuôn mặt: Theo dõi khuôn mặt trong video, cập nhật vị trí và kích thước của nó theo thời gian.
- Nhận dạng khuôn mặt: Nhận dạng người trong ảnh hoặc video dựa trên khuôn mặt của họ.
- Phát hiện tay: Xác định vị trí và cử chỉ của tay trong ảnh hoặc video.
- Xử lý ảnh: Thực hiện các thao tác xử lý ảnh cơ bản như xoay, thay đổi kích thước, cắt ghép và điều chỉnh độ sáng.
- Hiệu ứng video: Thêm hiệu ứng vào video như bộ lọc, chuyển tiếp và văn bản.

2.4.3. Ưu điểm

Dễ sử dụng: Cung cấp các lớp và hàm đơn giản, dễ hiểu.

Mã nguồn mở: Miễn phí cho mục đích thương mại và phi thương mại.

Cộng đồng tích cực: Cộng đồng người dùng và nhà phát triển hỗ trợ lẫn nhau.

Tích hợp với OpenCV: Sử dụng OpenCV để thực hiện các tác vụ xử lý ảnh và video nâng cao.

2.4.4. Nhược điểm

Thư viện mới: Còn đang được phát triển, một số tính năng chưa hoàn thiện.

Cộng đồng nhỏ hơn: Cộng đồng người dùng và nhà phát triển nhỏ hơn so với OpenCV.

2.4.5. Ứng dụng

Phát triển ứng dụng di động: Tạo ứng dụng di động có sử dụng camera như ứng dụng chụp ảnh, quay video, gọi điện video.

Phát triển trò chơi: Tạo trò chơi có sử dụng camera như trò chơi theo dõi chuyển động, trò chơi thực tế tăng cường.

Xử lý ảnh và video: Thực hiện các tác vụ xử lý ảnh và video như chỉnh sửa ảnh, cắt ghép video, thêm hiệu ứng.

Robot và tự động hóa: Điều khiển robot và hệ thống tự động bằng thị giác máy tính.

Y tế: Phân tích hình ảnh y tế, hỗ trợ chẩn đoán.

CHƯƠNG 3: ỨNG DỤNG GAME HỌC NGOẠI NGỮ

3.1. Thu thập dữ liệu từ vựng

- Dữ liệu từ vựng từ CEFR Level A1 đến C1 được thu thập từ danh sách Offord 5000 và 3000 words:

<https://www.oxfordlearnersdictionaries.com/wordlists/oxford3000-5000>

- Dữ liệu từ vựng CEFR Level C2 được thu thập từ <https://flyer.vn/wp-content/uploads/2022/10/Level-C2-Word-List.pdf>

- Phần dịch nghĩa được thu thập từ: <https://dictionary.cambridge.org/vi/>

Bảng thông tin số từ vựng được sử dụng theo cấp bậc của CEFR	
Bậc	Số từ trong dự án
A1	50
A2	50
B1	50
B2	50
C1	50
C2	50

3.2. Dò tìm bàn tay (Hand tracking) trong CV Zone

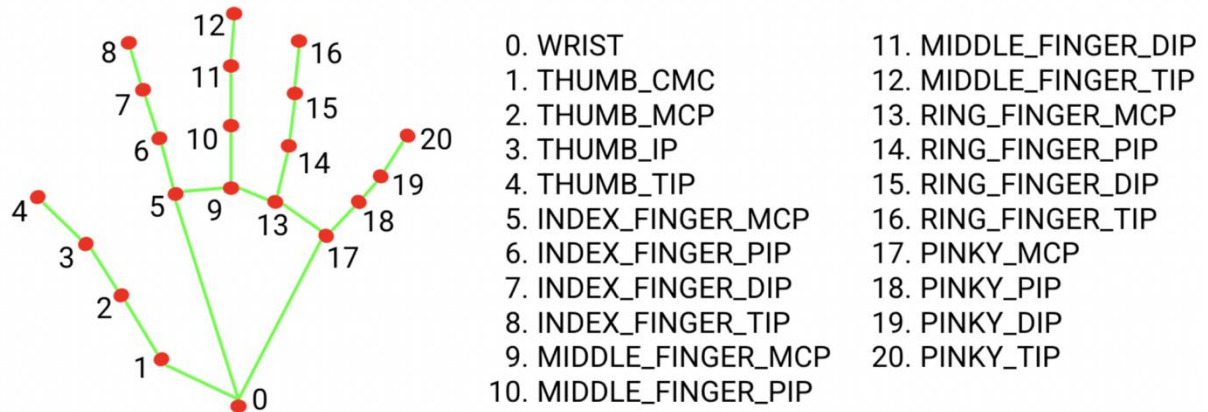
Thư viện CV Zone sử dụng mô hình Hand Landmarker của thư viện MediaPipe để nhận diện dữ liệu bàn tay từ hình ảnh.

Mô hình này là một gói mô hình bao gồm hai mô hình con, bao gồm:

+ Mô hình Palm Detection.

+ Mô hình Hand Landmarks Detection.

Mô hình định vị điểm mốc tay phát hiện vị trí chính xác của 21 tọa độ khớp tay trong các khu vực tay được phát hiện. Mô hình được huấn luyện trên khoảng 30.000 hình ảnh thể giới thực, cũng như một số mô hình tay tổng hợp được tạo ra trên nhiều nền khác nhau.



Mô hình Palm Detection xác định vị trí của các bàn tay trong hình ảnh đầu vào, và mô hình Hand Landmarks Detection xác định các điểm mốc tay cụ thể trên hình ảnh bàn tay được cắt ra do mô hình Palm Detection xác định.

Vì việc chạy mô hình Palm Detection mất thời gian, khi ở chế độ chạy video hoặc quay trực tiếp, Hand Landmarker sử dụng khung giới hạn vị trí bàn tay



được xác định bởi mô hình Hand Landmarks Detection trong một frame (khung hình) để xác định khu vực của bàn tay cho các khung hình sau. Hand Landmarker chỉ kích hoạt lại mô hình Palm Detection nếu mô hình Hand Landmarker không còn xác định

được sự hiện diện của bàn tay hoặc không thể theo dõi bàn tay trong khung hình. Điều này giảm số lần mô hình Hand Landmarker kích hoạt lại mô hình phát hiện Palm Detection.

3.3. Kết nối CV Zone với Unity

3.3.1. Gửi dữ liệu từ CV Zone

Cách thức mà chúng tôi sử dụng để truyền dữ liệu hand landmarks nhận được từ CV Zone cho Unity chính là thông qua giao thức UDP.

Trong dự án python handtracking đầu tiên chúng tôi tạo một đối tượng socket trong Python để tạo kết nối mạng, cụ thể là một socket sử dụng giao thức IPv4(socket.AF_INET) và giao thức UDP(socket.SOCK_DGRAM) để truyền dữ liệu:

Sau đó khai báo địa chỉ IP của máy chủ: `serverAddressPort = ("127.0.0.1", 5052)`

Phần tử đầu tiên là địa chỉ IP của máy chủ (trong trường hợp này là 127.0.0.1, là địa chỉ loopback, tức là máy chủ và máy khách đều đang chạy trên cùng một máy tính).

Phần tử thứ hai là cổng mà máy chủ đang lắng nghe để nhận dữ liệu. Trong trường hợp này, cổng là 5052.

Cuối cùng, sau khi đã xử lý dữ liệu để có thể sử dụng được trong dự án Unity, chúng tôi gửi dữ liệu từ một socket UDP đến địa chỉ IP và cổng được chỉ định:

```
sock.sendto(str.encode(str(data)), serverAddressPort)
```

3.3.2. Nhận dữ liệu trong Unity

Để có thể nhận được dữ liệu thông qua giao thức UDP trong Unity chúng tôi tạo các thành phần cần thiết bao gồm: luồng nhận dữ liệu, đối tượng UDP client nhận dữ liệu, cổng để UDP client lắng nghe, biến cờ cho biết tín hiệu đã bắt đầu hoặc ngừng nhận dữ liệu và cuối cùng là biến chứa dữ liệu.

- Khởi tạo một luồng mới để nhận dữ liệu:

```
receiveThread = new Thread(new ThreadStart(ReceiveData));
```

```
receiveThread.IsBackground = true; // Đánh dấu luồng là nền
```

```
receiveThread.Start(); // Bắt đầu luồng
```

- Phương thức ReceiveData() để nhận dữ liệu:

Khởi tạo UDP client để nhận dữ liệu từ cổng đã chỉ định

```
udpClient = new UdpClient(port);
```

```
IPEndPoint anyIP = new IPEndPoint(IPAddress.Any, 0);
```

```
// Nhận dữ liệu từ máy chủ và lưu vào mảng bytes
```

```
byte[] dataByte = udpClient.Receive(ref anyIP);
```

3.4. Ứng dụng Game học ngoại ngữ

Giới thiệu: DLU Adventure là một ứng dụng trò chơi dành cho việc học ngoại ngữ, chủ yếu là từ vựng. Mục tiêu của ứng dụng là cung cấp một trải nghiệm học tập thú vị và tương tác để người dùng có thể cải thiện khả năng ngôn ngữ của mình một cách hiệu quả.

Cách chơi: người chơi vẫn sẽ sử dụng chuột để di chuyển qua các phân cảnh của trò chơi và khi đến phân cảnh chính của trò chơi trắc nghiệm, khi người chơi đưa tay vào phạm vi quan sát của camera thì sẽ thấy một bàn tay 3D có hình dáng giống với bàn tay của mình. Người chơi sẽ di chuyển bàn tay để nhấn các nút bấm 3D để chọn câu trả lời cho các câu hỏi trắc nghiệm.

Cách thức hoạt động của trò chơi:

Sau khi trò chơi nhận được các dữ liệu về tọa độ các điểm trên bàn tay của người chơi, trò chơi sẽ tạo ra một bàn tay 3D trông giống với tay của người chơi, khi người chơi di chuyển tay của mình trong phạm vi quan sát của camera thì bàn tay 3D

cũng sẽ di chuyển theo điều này cho phép người chơi có thể tương tác với các vật thể 3D trong trò chơi.

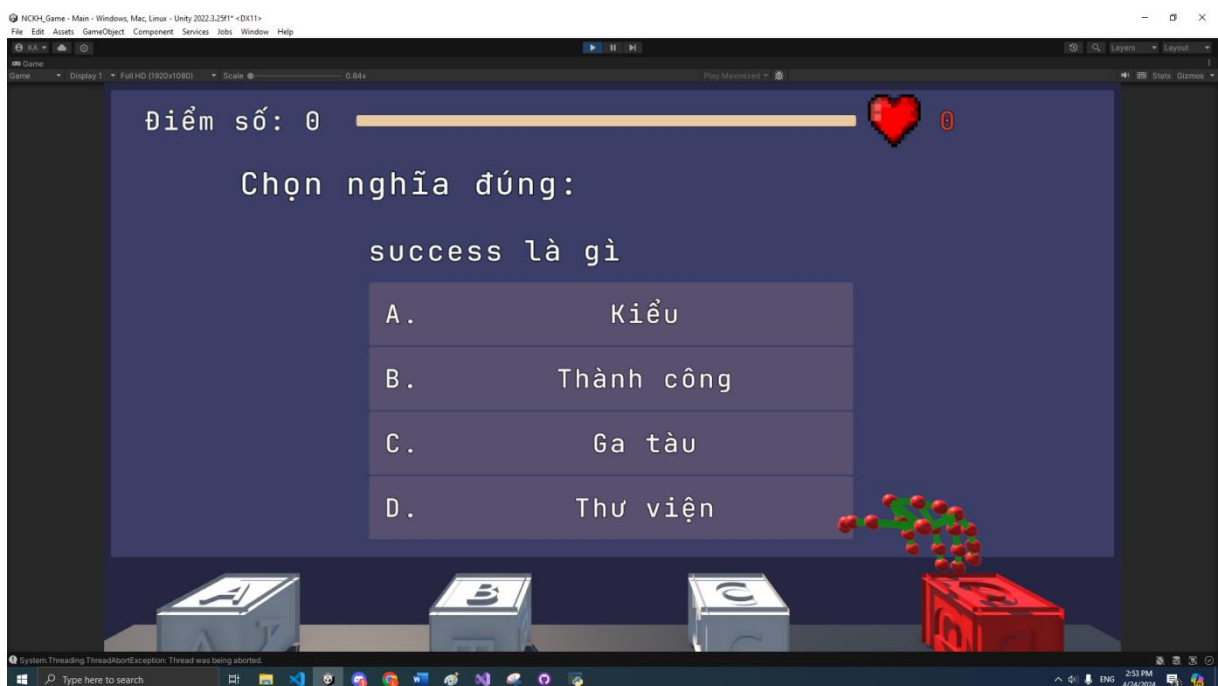
3.5. Một số hình ảnh sản phẩm



MainMenu (Trang chủ): nơi bắt đầu của trò chơi gồm các nút start để bắt đầu chơi, nút setting để chỉnh các tùy chỉnh như âm thanh, nút About để xem thông tin người làm trò chơi, nút Quit để thoát trò chơi.



Select Level (Chọn cấp độ): chọn các cấp độ từ vựng tiếng anh CEFR từ A1 đến C2.



Main: màn hình chơi trắc nghiệm, người chơi sẽ chọn 4 lựa chọn để trả lời bằng cách dùng tay thực hành động tác ấn nút hoặc chạm vào nút 3D trước camera.

KẾT LUẬN

Kết luận từ việc kết hợp thư viện CVZone và UnityEngine trong việc xây dựng mô hình trò chơi, nhóm đã chứng minh được tiềm năng và khả năng ứng dụng rộng rãi của công nghệ trí tuệ nhân tạo (AI) trong lĩnh vực phát triển trò chơi và phần mềm bổ ích. AI đã và đang đóng vai trò quan trọng trong việc cải thiện chất lượng trò chơi, tạo ra trải nghiệm chơi game thực tế và thú vị hơn và tạo ra nội dung động và đa dạng.

CVZone, một gói phần mềm trí tuệ nhân tạo dựa trên OpenCV và Mediapipe, giúp đơn giản hóa quá trình xử lý hình ảnh và các chức năng AI, tạo ra cơ hội mở rộng cho việc tạo ra các trò chơi và phần mềm bổ ích phức tạp hơn.

Tóm lại, việc kết hợp thư viện CVZone và UnityEngine đã chứng minh sự tiềm năng và khả năng ứng dụng rộng rãi của AI trong việc xây dựng trò chơi và phần mềm bổ ích, tạo ra trải nghiệm chơi game thực tế và thú vị hơn, cũng như tạo ra nội dung game sống động và đa dạng hơn.

DANH MỤC TÀI LIỆU THAM KHẢO

- CAMBRIDGE. (n.d.). *cambridgeenglish*. Retrieved from English Language Assessment: <https://www.cambridgeenglish.org/vn/exams-and-tests/cefr/>
- Google. (n.d.). *Mediapipe*. Retrieved from Google for Developers: <https://developers.google.com/mediapipe>
- Kaplan, A. (2021). *Artificial Intelligence, Business and Civilization*. January 14, 2022 by Routledge.
- OpenCV. (n.d.). *About*. Retrieved from OpenCV University: <https://opencv.org/about/>
- Victor Wiley, T. L. (2018). Computer Vision and Image Processing. In T. L. Victor Wiley, *International Journal Of Artificial Intelligence Research* (p. 28). Jakarta Indonesia: STMIK Dharma Wacana.